Towards an Indian Land Data Assimilation System (ILDAS): A coupled hydrologic-hydraulic system for water balance assessments

Bhanu Magotra[a], Ved Prakash[a], Manabendra Saharia[a,b], Augusto Getirana[c,d], Sujay Kumar[c], Rohit Pradhan[e], C.T. Dhanya[a], Balaji Rajagopalan[f], Raghavendra P. Singh[g], Ayush Pandey[a], Mrutyunjay Mohapatra[h]

[a] Department of Civil Engineering, Indian Institute of Technology Delhi, Hauz Khas, New Delhi 110016, India

[b] Yardi School of Artificial Intelligence, Indian Institute of Technology Delhi, Hauz Khas, New Delhi 110016, India

[c] Hydrological Sciences Laboratory, NASA Goddard Space Flight Center, Greenbelt, MD, USA

[d] Science Applications International Corporation, Greenbelt, MD, United States

[e] Indian Institute of Remote Sensing, Indian Space Research Organization, Dehradun, 248001, India

[f] Department of Civil, Environmental and Architectural Engineering & CIRES, University of Colorado, Boulder CO, USA

[g] Space Applications Centre, Indian Space Research Organization, Ahmedabad 380015, India

[h] Indian Meteorological Department, New Delhi 110003, India

Corresponding Author:

Manabendra Saharia

Indian Institute of Technology Delhi

New Delhi, India 110016

Office: +91-011-26591260

Email: msaharia@iitd.ac.in

ABSTRACT

Effective management of water resources requires reliable estimates of land surface states and fluxes, including water balance components. But most land surface models run in uncoupled mode and do not produce river discharge at catchment scales to be useful for water resources management applications. Such integrated systems are also rare over India where hydrometeorological extremes have wreaked havoc on the economy and people. So, an Indian Land Data Assimilation System (ILDAS) with a coupled land surface and a hydrodynamic model has been developed and driven by multiple meteorological forcings (0.1°, daily) to estimate land surface states, channel discharge, and floodplain inundation. ILDAS benefits from an integrated framework as well as the largest suite of observation records collected over India and has been used to produce a reanalysis product for 1981-2021 using four forcing datasets, namely, Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2), Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS), ECMWF's ERA-5, and Indian Meteorological Department (IMD) gridded precipitation. We assessed the uncertainty and bias in these precipitation datasets and validated all major components of the terrestrial water balance, i.e., surface runoff, soil moisture, terrestrial water storage anomalies, evapotranspiration, and streamflow, against a combination of satellite and in situ observation datasets. Our assessment shows that ILDAS can represent the hydrological processes reasonably well over the Indian landmass with IMD precipitation showing the best relative performance. Evaluation against ESA-CCI soil moisture shows that MERRA-2 based estimates outperform the others, whereas ERA-5 performs best in simulating evapotranspiration when evaluated against MODIS ET. Evaluations against observed records show that CHIRPS-based estimates have the highest performance in reconstructing surface runoff and streamflow. Once operational, this system will be useful for supporting transboundary water management decision making in the region.

Keywords: Indian Land Data Assimilation System (ILDAS), water balance assessments, streamflow, south Asia

1. Introduction

Effective water resources management requires consistent and long-term estimates of the terrestrial water balance, usually derived from computational models driven by accurate meteorological forcings and observational inputs. Land surface models (LSMs) are used to mathematically model the various land surface processes critical in transferring energy fluxes and moisture between the land surface and the atmosphere. The primary purpose of an LSM is to simulate the dynamics of water storage, energy, and water fluxes on the surface and subsurface, by using physically based equations (Kirchner, 2006). While LSMs have been used in multiple studies to simulate water balance at a continental scale across the world and over India, an integrated hydrologic-hydraulic system over the Indian subcontinent has not been developed.

The Indian mainland consists of complex terrain with a surface elevation ranging from approximately 10 to 8000 meters above mean sea level while having distinct topography that includes eastern and western coastal regions, northern and northeastern mountain ranges, central flood plains, southern peninsula, and western arid regions. Moreover, the Indian climate is quite diverse, with annual mean temperature and precipitation ranging from approximately 7 to 27 ºC and 500 to 4900 mm, respectively. Being primarily an agrarian economy, India relies heavily on the long-term and seasonal availability of freshwater. Additionally, many regions of India are often exposed to natural hazards such as floods and droughts, which are associated with intense precipitation during the southwest monsoon and hot and dry summers, respectively (Saharia et al., 2021; Zhang et al., 2017). Moreover, the warming climate has increased the uncertainty in precipitation, further exaggerating the risks associated with short-term and long-term variations in the natural water balance (Ali and Mishra, 2018). Therefore, accurate estimates of land surface states, streamflow, and flood plain inundation are critical in the decision-making process to ensure national food security, natural hazards mitigation, and water resources planning and management. However, to generate these estimates, two primary challenges need to be addressed: (a) the representation of the spatial variability of various land surface processes and the initial states in LSMs for such a complex landmass is difficult (Zhao & Li, 2015), and (b) the models run in a non-operational setting where the LSMs are generally not coupled with a routing model, and thus, lack the ability to provide near real-time estimates of streamflow at catchment scales. To address this, we set up an Indian Land Data Assimilation System (ILDAS), which is based on a land surface and hydrodynamic model coupled in an offline mode (i.e., no feedback between LSM and hydrodynamic model) and is driven by multiple meteorological forcings to generate spatially consistent and high-resolution estimates of land surface states, water balance, and energy fluxes over the Indian mainland.

A Land Data Assimilation System (LDAS) facilitates the assimilation of in situ observations and remotely sensed data to improve the accuracy of LSMs through various data assimilation techniques and the use of observation-based atmospheric forcing data (Kumar et al., 2014). The progress towards the development of various LDAS was led by the North American LDAS (NALDAS; Lohmann et al., 2004) and Global LDAS (GLDAS; Rodell et al., 2004), which were initially developed to provide optimal land surface states and fluxes to atmospheric models to improve weather and climate predictions (Xia et al., 2019). With the increasing availability of remotely sensed data, enhanced in situ observation gauge networks and affordable computational power, many regional LDAS have been developed, such as European LDAS (ELDAS; (Jacobs et al., 2008), South American LDAS (SALDAS; (de Goncalves et al., 2006), South Asia LDAS (Ghatak et al., 2018), and Canadian LDAS (CaLDAS; Carrera et al., 2015). The water and energy fluxes, along with other land surface states generated by regional and global LDAS, have found wide usability in various applications such as flood and drought monitoring, climate prediction models, water resource management, and agricultural crop management (Jin et al., 2018; McNally et al., 2017; Sawada and Koike, 2016; Yucel et al., 2015). The ILDAS is built on NASA's Land

Information System Framework (LISF; lis.gsfc.nasa.gov), which is an open-source software that enables a multi-model, multi-data approach to land surface modeling (Kumar et al., 2006). As part of a series of studies that will be carried out towards establishing ILDAS, this paper presents the results from the first study in which we used the Noah land surface model with multiparameterization options (Noah-MP; Niu et al., 2011) coupled with the Hydrological Modeling and Analysis Platform (HyMAP; Getirana et al., 2017, 2012) to simulate hydrological processes over the Indian landmass using multiple global meteorological forcing datasets, namely, Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2; Gelaro et al., 2017), Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS; Funk et al., 2015), ECMWF's ERA-5 (ERA-5; Hersbach et al., 2020), and IMD's gridded precipitation over India. The previous studies over India have mainly used LSMs without a coupled hydrodynamic model and are focused primarily on better representation and understanding of various processes involved in energy and water cycle (Attada et al., 2018; Ghodichore et al., 2022; Maity et al., 2017; Nair & Indu, 2019; Patil et al., 2011). In the study conducted using South Asia LDAS (Ghatak et al., 2018), the authors focus on effects of precipitation uncertainty on various hydrological simulations including streamflow over a similar spatial domain as that of ILDAS, but it is limited to a relatively short period of evaluation and fewer observed streamflow locations. Moreover, the study does not include streamflow evaluation over India's geographical domain. In this study, we present a comprehensive evaluation of surface runoff, soil moisture, terrestrial water storage anomalies, evapotranspiration, and streamflow. Besides evaluating major components of the water balance, we also assessed the uncertainty and bias due to spatiotemporal heterogeneity in the forcing precipitation by evaluating against the gauge-based gridded precipitation provided by the Indian Meteorological Department (IMD).

Overall, the objectives of this study are to:

(a) set up ILDAS by coupling a land surface and hydrodynamic model to generate a high-resolution reanalysis dataset over the Indian domain.

(b) quantify the uncertainty and bias in precipitation provided by the global forcings over the Indian mainland.

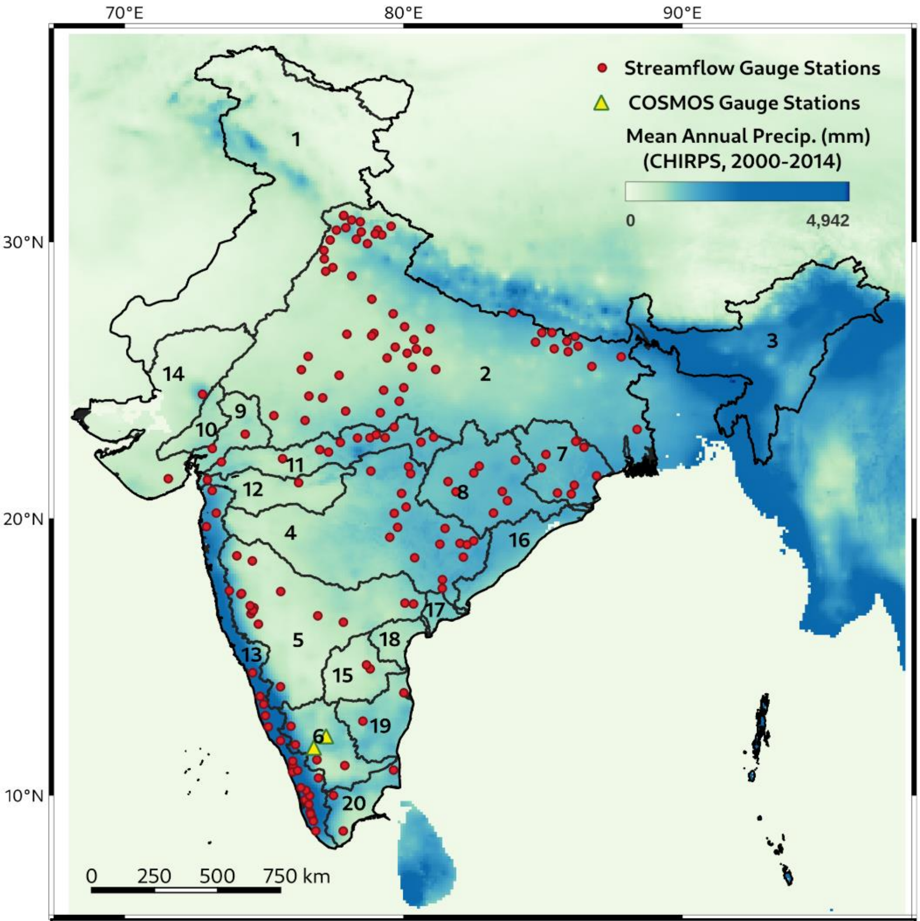
(c) evaluate ILDAS nationwide performance by evaluating simulated water balance variables against in situ and satellite-observed products.

The paper is organized as follows: section 2 describes the study area and various datasets used in this study. It also briefly explains the Noah-MP model and the methodology involved in running the model and evaluation of results. In section 3, results are presented along with relevant discussion. Finally, section 4 provides the conclusions of the study and future work.

2. Data and methods

2.1 Study Area

The modeling system is defined on a spatial domain spanning 68°E – 98°E and 5.5°N – 37.5°N, as shown in (Fig. 1). The landmass primarily consists of the geographical region of India along with some portions of neighboring countries. By taking a wider geographical extent than India's political boundary, we ensured that the LSM could process the necessary meteorological and geological information at the boundary of the Indian landmass. India has a diverse climate and geography that can be attributed to being the world's seventh largest country in terms of area. The Indian mainland includes mountain ranges in the north and north-east, the western and eastern coastal regions, the Indo-Gangetic plains, the desert in western Rajasthan, the peninsular plateau and the islands of Lakshadweep and Andaman and Nicobar. The overall climate of India is considered tropical, with a mixture of dry and wet tropical weather in the country's interior regions. The country gets most of its precipitation from monsoon rains that begin in June and last till September. Although the analysis has been done on a 0.1° spatial resolution grid across the Indian landmass, an attempt has been made to highlight the outcomes based on major river basins as specified by the Central Water Commission (CWC), India, which are available through the India Water Resources Information System (IWRIS; www.india-wris.nrsc.gov.in).



List of Major River Basins (CWC,India)		
1 Indus (Upto Border)	7 Brahmani & Subernarekha	13 West-Flowing Rivers (Tapi-Kanyakumari) (WFR-TK)
2 Ganga	8 Mahanadi	14 West-Flowing Rivers (Kutch and Saurashtra incl. Luni) (WFR-KSL)
3 Brahmaputra, Barak & Others	9 Mahi	15 Pennar
4 Godavari	10 Sabarmati	16-18 East-Flowing Rivers (Mahanadi-Pennar) (EFR-MP)
5 Krishna	11 Narmada	19-20 East Flowing Rivers (Pennar-Kanyakumari) (EFR-PK)
6 Cauvery	12 Tapi	

Fig. 1. A map depicting the ILDAS spatial domain, Indian Central Water Commission River basins and streamflow gauge stations considered in the study.

2.2 Modeling framework

LISF is an infrastructure that supports multiple land surface models, meteorological forcings, and various data assimilation and routing schemes. Given the scalability and flexibility of LISF, it is well suited for large-scale terrestrial modeling as it enables users to harness high-performance computing and combine various modeling tools and data sources in a systematic and streamlined manner. The Noah land surface model with multiparameterization options (Noah-MP; (G. Y. Niu et al., 2011) builds upon the earlier Noah model (Ek et al., 2003) by including newer land surface physics such as (a) tiling scheme in the grid, which can differentiate between vegetation and bare soil, (b) a multi-layer snowpack as compared to one bulk-layer snowpack, (c) a canopy layer, (d) separation of permeable and non-permeable frozen soil fractions, and (e) TOPMODEL-based runoff scheme along with Simple Groundwater Model (SIMGM; Niu et al., 2007). The Noah-MP also includes multiparameterization options for various physical processes such as runoff generation, dynamic vegetation, canopy stomatal resistance, groundwater, and so on.

To simulate discharge and floodplain inundation, we use the coupled Noah-MP with the Hydrological Modeling and Analysis Platform (HyMAP; (A. Getirana, Peters-Lidard, et al., 2017; A. C. V. Getirana et al., 2012) river routing model. HyMAP is a state-of-art global scale hydrodynamic model that uses local inertia formulation to simulate surface water dynamics in rivers and floodplains based on baseflow and surface runoff provided by the LSM at each modeling timestep (Bates et al., 2010; De Almeida et al., 2012). The model employs the local inertia formulation, which involves solving the complete momentum equation of open channel flow. This enables a more stable and efficient representation of river flow diffusiveness and inertia of large water masses with deep flow. Such a representation is important for a physically accurate representation of wetlands, lakes, floodplains, tidal effects, and impoundments (A. Getirana et al., 2020). It adopts a sub-grid approach where both base flow and surface runoff at each grid cell are passed through individual linear reservoirs and adjusted against relevant time delay factors. To derive water storage, elevation and discharge in stream and floodplains, HyMAP processes the topographic information in the form of Digital Elevation Model (DEM), river geometry, and roughness. The HyMAP parameters are derived from the Multi-Error-Removed Improved-Terrain (MERIT; Yamazaki et al., 2017) DEM while the widths of major rivers are derived from MERIT-Hydro which is a 90-m global estimated river width dataset based on Landsat data. However, the width of smaller channels that were not detected by the dataset, was derived using an empirical equation (A. C. V. Getirana et al., 2012):

$$w = \max\left(0.2\,,\ 20 \times Q_{med}^{0.5}\right) \quad (1)$$

where w (m) is the average river width within a grid cell and Q_{med} (m^3/s) is the annual mean discharge.

River width and bankfull height, h (m) was estimated using the following empirical equation:

$$h = \max(0.35, \alpha \times w)\ ;\ where\ \alpha = 2.6 \times 10^{-3} \quad (2)$$

The roughness of open channels as well as floodplains is considered in the form of Manning's coefficient, which is based on vegetation type in the individual grid cell (A. C. V. Getirana et al., 2012).

Model Configuration:

The specifications of various ILDAS parameters are shown in Table 1. The Land Data Toolkit (LDT; Arsenault et al., 2018) was used to generate parameter files that contain various static information to be processed by Noah-MP, such as land use/land cover, irrigation, soil types, elevation, and so on. Four open-loop individual runs were conducted in retrospective mode within the LIS Framework (LISF) on a 0.1x0.1 grid at a 15 minutes timestep. The four runs were conducted from 1981-2021 using MERRA-2, CHIRPS, ERA-5, and IMD, respectively, and the model outputs were produced at daily timestep. In our initial testing, we found that the model reached an equilibrium state over the ILDAS domain after approximately ten years of simulation. We evaluated the equilibrium of the model based on the percentage difference between the water balance components generated over two consecutive spin-up runs (Rodell et al., 2005). To ensure that the model has significant atmospheric information to reach a steady state, we performed two spin-ups for each run with five years of meteorological data. The simulations were performed in a high-performance computing facility using 64-100 total CPUs with an average completion time of approximately 3 hours per year of simulation.

Table 1. List of ILDAS components and their specifications.

ILDAS Component	Specifications
Land Surface Model	Noah-MP 3.6
Routing Scheme	HyMAP
Spatial Extent	68°-98°E, 5.5°-37.5°N
Spatial Resolution	0.1°
Temporal Resolution	15 minutes Noah-MP 3.6 and HyMAP with adaptive timestep, daily output fields
Time Period	1981-2021
Forcing	MERRA-2, CHIRPS (precipitation) + MERRA-2, ERA-5, IMD
Forcing Variables	Precipitation, near-surface air temperature, near-surface specific humidity, surface pressure, eastward and northward wind velocity, incident longwave and shortwave radiation
Forcing Height	2 m for surface air temperature, specific humidity, and surface pressure, 10 m for wind
Topography and river network	MERIT Hydro
Soils Definition	(NCAR) STATSGO+FAO blended soil texture map
Vegetation Definition	MODIS-IGBP (NCEP-modified), Monfreda et al. (2008) crop types
Output Format	NetCDF

2.3 Atmospheric Forcings

2.3.1 MERRA-2

The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2; Gelaro et al., 2017) improves upon its predecessor, MERRA, by leveraging recent developments at NASA's Global Modeling and Assimilation Office (GMAO), which include updates to the Goddard Earth Observing System (GEOS) as well as new assimilation schemes for microwave observations, NASA ozone observations, hyperspectral radiance and many more datatypes (Gelaro et al., 2017). In previous studies (Ghatak et al., 2017, 2018; Gupta et al., 2020), MERRA-2 has shown satisfactory results for temperature and precipitation estimates over India. In this study, we used bias-corrected precipitation from the MERRA-2 dataset at a spatial resolution of 0.625°x0.5° and hourly timesteps for the period 1981-2021.

2.3.2 CHIRPS

Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS; Funk et al., 2015) is a quasi- global precipitation dataset derived from global Cold Cloud Duration (CCD) rainfall estimates calibrated using Tropical Rainfall Measuring Mission Multi-Satellite Precipitation Analysis version 7 (TMPA 3B42 v7). CHIRPS aims to bridge the gap between high latency precipitation products such as Global Precipitation Climatology Centre (GPCC) and low latency satellite-only products like the TMPA 3B42 RT (Funk et al., 2015). The CHIRPS precipitation estimates incorporate in situ gauge station data and active radar satellite systems, and the dataset is available from 1981 to the near present at a high spatial resolution of 0.05°. Since CHIRPS consists of only precipitation, we used MERRA-2 as an overlay to provide other variables in the simulation.

2.3.3 ERA-5

ECMWF's ERA-5 (ERA-5; Hersbach et al., 2020) is a new global reanalysis dataset that builds upon the earlier ERA-Interim reanalysis. The dataset is available from 1959 to present at a spatial resolution of 0.25° and hourly timestep. ERA-5 uses a 10-member ensemble with 12hr 4D-Var data assimilation method to include various reprocessed datasets. The dataset is available in preliminary form at 5 days latency to real time and the final quality assured product is released with a latency of 2 months.

2.3.4 Indian Meteorological Department (IMD) Precipitation

We used gridded daily rainfall data related to the 1981-2021 period at 0.25° spatial resolution provided by the Indian Meteorological Department (IMD). This data is generated using 6955 gauge stations which include IMD observatory stations, hydrometeorological observatories, and Agromet observatories (Pai et al., 2014). To generate gridded data from point-based station rainfall data, an inverse distance weighted interpolation scheme with localized directional effects and barriers was used which is based on (Shepard, 1968). While comparing the observed and reanalysis precipitation products, we rescaled all the datasets to a 0.1° spatial resolution. Since IMD precipitation only covers the geographical boundaries of India, we supplemented it with MERRA-2 to provide the missing values beyond IMD precipitation's domain.

2.4. Observed and Satellite Products

The following section covers the suite of satellite-based and in-situ observations that were acquired from various sources for the evaluation.

2.4.1 GRUN (Runoff)

We used the GRUN runoff dataset (Ghiggi et al., 2019) as observed data, which is an observationally driven growly reconstructed monthly runoff at 0.5° resolution for the period of 1902-2014. Machine learning-based Random Forest (RF) was used to generate the GRUN runoff data, and the temperature and precipitation gridded data was used from the Global Soil

Wetness Project Phase 3 (GSWP3) (Kim & Office, 2017). The Global Streamflow Indices and Metadata archive (GSIM) was used to obtain monthly runoff observations. To test the sensitivity of the machine learning algorithm, Ghiggi et al., (2019) used 50 ensembles. In this study, we used the GRUN reconstruction, which is an ensemble mean of the realizations.

2.4.2 Observed Soil Moisture (COSMOS)

We acquired in-situ observed soil moisture at two-gauge stations (Singanallur-SGR and Adahalli-MDH) from Indian Cosmic Ray Network (ICON; Upadhyaya et al., 2021). ICON consists of seven sites equipped with COSMOS instruments across India operational from 05/2015. The Cosmic Ray Neutron Probe (CRNP) technique is used in the COSMOS instrument, which uses non-invasive neutron counts as a measure of soil moisture. More information can get from webpage https://cosmos-india.org/ .

2.4.3 ESA-CCI (Satellite Soil Moisture)

European space agency's climate change initiative for soil moisture (ESA-CCI SM version v07.1) is used as the soil moisture reference dataset in our study, which is available at 0.25° resolution from 1978 (Dorigo et al., 2017; Gruber et al., 2017, 2019). The ESA CCI provides three products, namely Active, Passive, and Combined. While Active products were retrieved from active microwave sensors using the TU Wien water Retrieval Package (WARP) algorithm, their Passive counterparts were obtained using the Land Parameter Retrieval Model (LPRM) algorithm (Owe et al., 2008) from passive-microwave-based sensors. In this study, we used a combined product of the ESA-CCI SM products, which is generated by merging the active and passive products following a decision tree method and has been found to be most suitable for evaluation in the Indian mainland (Chakravorty et al., 2016; Maina et al., 2022).

2.4.4 MODIS Evapotranspiration

We acquired Evapotranspiration from Moderate Resolution Imaging Spectroradiometer (MODIS). We used the MYD16A2GF product (Running et al., 2019), a gap filled at 8-day temporal resolution and 500m spatial resolution. Calculation of ET is typically based on the conservation of either energy or mass or both. The Penman-Monteith equation (J. L. Monteith, 1965) has been used in the ET algorithm.

2.4.5 GRACE Terrestrial Water Storage Anomalies (TWSA)

We acquired Gravity Recovery Climate Experiment (GRACE; Landerer and Swenson, 2012; Tellus, 2018) terrestrial water storage anomalies (TWSA) data from the Jet Propulsion Laboratory (JPL). The TWS is obtained using the Mass Concentration blocks (mascons) techniques, which implement geophysical constraints referred to as mascons. Our study used the latest JPL mascon solution (Tellus, 2018; Watkins et al., 2015), which is available monthly at 0.5° spatial resolution. The anomalies are calculated relative to the January 2004-December 2009 as the time-mean baseline and provided as TWSA.

2.4.6 Observed streamflow

The daily streamflow observed records were collected from various government agencies through the public domain as well as official requests. The records were checked for data inconsistencies and were converted to a standard format for analysis. Due to varying record lengths across the gauge stations, a common testing period with enough stations could not be established. We selected gauge stations with at least twenty years of daily recorded values in the period 1981-2021, which may or may not be continuous. In this way, a total of 162 stream flow gauge stations were chosen across the study area, as highlighted in (Fig. 1).

2.5 Mann-Kendall Trend Analysis

We used the Mann-Kendall (Mann, 1945) test for trend analysis, which is a non-parametric test for the monotonic trends of environmental data over time, such as climate data or hydrological data (Hu et al., 2020). It is a rank-based significance test, that identifies the significance of the trend by checking S-statistics of the time series fall in the confidence interval null hypothesis or not. The S-statistics are calculated to determine whether the trend is increasing or decreasing.

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} sign(x_j - x_k) \tag{3}$$

where x is the time series variable, and the subscripts j and k are the observation time. $sign(x_j - x_k)$ is equal to +1, 0, or -1, which means increasing, no, and decreasing trends, respectively. We rescaled the S-statistics between (-1,1) for better understanding. Here we assume that there is no significant trend in data at a level of 5% (or 95% confidence interval) as a null hypothesis.

2.6 Evaluation criteria

To check the effectiveness of different meteorological forcing, we used correlation coefficient, Relative Root Mean Square Error (RRMSE), and percent bias to evaluate the annual mean precipitation of reanalyzed meteorological forcings with gridded precipitation data from IMD.

$$RRMSE = \frac{\sqrt{\frac{\sum (p_o - p_r)^2}{N}}}{\frac{\sum p_o}{N}} \tag{4}$$

$$Pbias = \frac{\sum (p_r - p_o)}{\sum p_o} \times 100\% \tag{5}$$

where p_o, p_r, and N are Observed, reanalyzed and number of data, respectively.

To perform a balanced assessment of simulated water balance against observed values, we rescaled all the observed datasets to the same resolution of ILDAS for the evaluation of the

ILDAS. We selected Kling Gupta Efficiency (KGE; Gupta et al., 2009) as our primary metric with its three components, namely, correlation coefficient (*r*), variability ratio(α) and bias(β). The calculation of KGE is expressed as:

$$KGE = 1 - \sqrt{S_r[r-1]^2 + S_\alpha[\alpha-1]^2 + S_\beta[\beta-1]^2} \qquad (6)$$

$$\left(\alpha = \frac{\sigma_s}{\sigma_o}, \beta = \frac{\mu_s}{\mu_o}\right)$$

where Sr, Sα, and Sβ are scaling factors for the three components respectively, that can be specified by the user; σs and σo are the standard deviations for simulated and observed variables, respectively, and μs and μo are the corresponding mean values. The three components of KGE highlight different parts of the performance of a model where the agreement between the timing of simulated and observed values is given by correlation (r), the statistical variability is expressed by variability ratio (α), and the bias is highlighted by bias (β). A KGE value equal to 1 (r=1, α=1, β=1) means a perfect agreement between simulated and observed values, while a value less than -0.41 denotes that the model is a worse predictor than the mean of the observed series (Knoben et al., 2019). The scaling factors can be used to emphasize one or more components of KGE depending on the objective of the study (Mizukami et al., 2019). In this study, we wanted to present a balanced overview of performance, and therefore, we considered all three scaling factors equal to 1.0. Moreover, considering the wide differences in soil moisture obtained from models, satellites, and in-situ observations, we used anomaly correlation and unbiased RMSE (ubRMSE) instead of KGE for an objective evaluation of the variable.

3. Results and discussion

3.1 Precipitation analysis

A significant uncertainty in hydrological models comes from meteorological forcings, particularly precipitation. In particular, the precipitation frequency distribution is the most important factor for the accurate characterization of frequent and extreme floods (Newman et al., 2021). To check consistency, we evaluated the meteorological forcing inputs from 1981 to 2021 against the IMD gridded observed precipitation dataset. Overall, ERA-5 shows a better national median correlation (median value of the correlation for all gridded values in the Indian mainland) compared to MERRA-2 and CHIRPS. It also shows a better correlation with IMD in the Himalayan and northeast regions than MERRA-2 and CHIRPS (Fig. 2a-c). However, MERRA-2 shows a better correlation in Rajasthan and Deccan plateau than CHIRPS and ERA-5 meteorological forcing precipitation. In all three meteorological forcings, we found that precipitation is underestimated (i.e., percent bias shows negative values) in the Western Ghats, Himalayan, and northeastern region (Fig. 2d-f). General underestimation in satellite precipitation over the Himalayan region has also been reported by Bharti and Singh, (2015) due to satellites missing the convective clouds. CHIRPS indicates positive Pbias (overestimation) in parts of northeast India. In the northern plane and Deccan plateau, CHIRPS and ERA-5 show overestimation (positive Pbias) compared to MERRA-2. Moreover, CHIRPS and ERA-5 show a nationwide median of Pbias positive (overestimation), whereas MERRA-2 shows underestimation. Underestimation in the Western Ghats might be due to the radiometrically warm land surface, and the coastal regions are a mixture of the radiometrically cold oceans (Mccollum & Ferraro, 2005; Shah & Mishra, 2016). We also found that RRMSE is higher in the Western Ghats, Himalayan, and northeast regions than in the Deccan plateau and Rajasthan (semiarid areas) (Fig. 2g-i). Moreover, all metrological forcings show improvement in RRMSE for the monsoon (JJAS) period (Fig. S4). Fig. 3a shows the annual mean precipitation for IMD, MERRA-2, CHIRPS, and ERA-5. We identified that MERRA2 underestimates the annual mean precipitation compared to other forcings till 2009, which is consistent with the observations of a previous study

(Bhattacharyya et al., 2022). From the figure, it is clear that all meteorological forcings are showing a significant positive trend ($p<0.05$) unlike IMD for most of the basins. MERRA-2 displayed an increasing trend in the Himalayan regions, a pattern similar to what was found by Yoon et al., (2019). Overall, our analysis shows that CHIRPS is less uncertain than MERRA-2 and ERA-5 with IMD as the baseline.

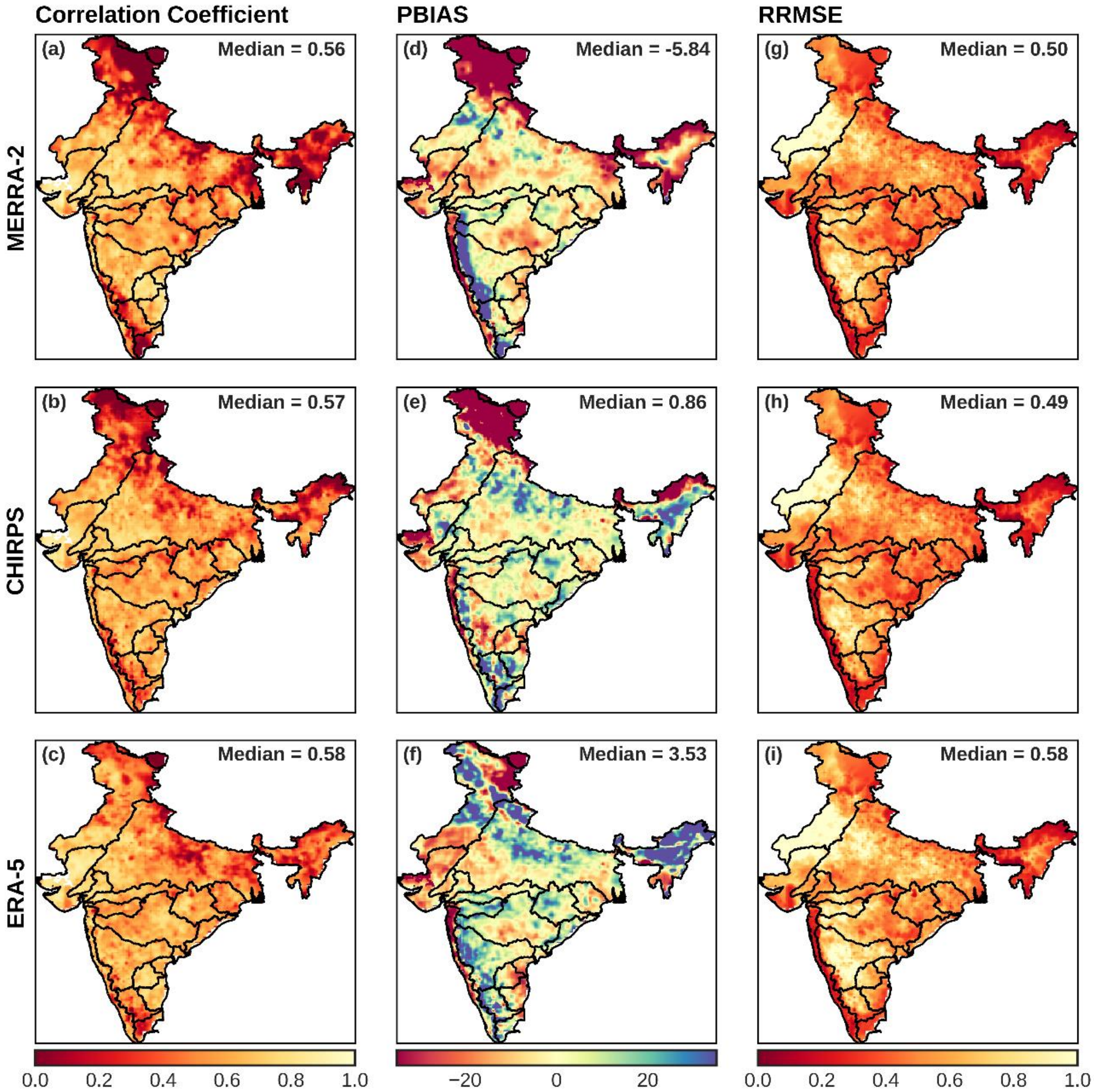


Fig. 2. Comparison of correlation coefficient (a-c), Pbias (d-f), and RRMSE (g-i) of Annual precipitation for different precipitation datasets (i.e., MERRA-2, CHIRPS, and ERA-

5) with respect to the India Meteorological Department (IMD) precipitation dataset for 1981-2021.

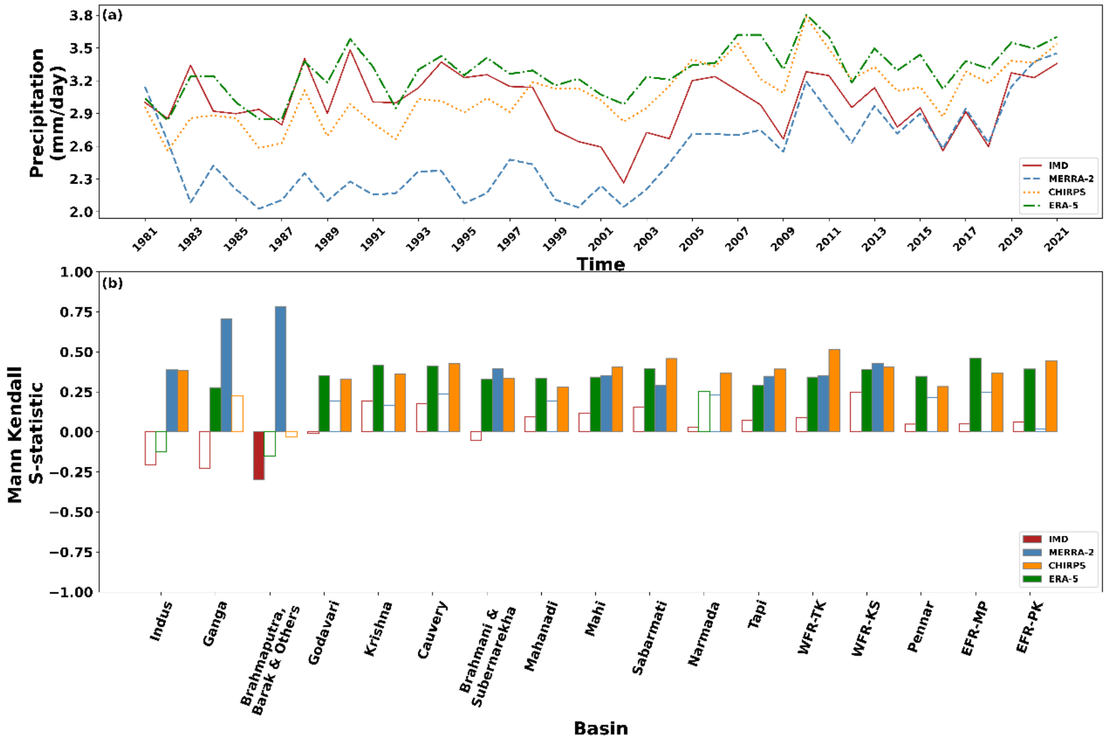


Fig. 3. Nationwide mean annual precipitation plot for different precipitation datasets (IMD, MERRA-2, CHIRPS, ERA-5) for 1981-2021(a) and basin wise precipitation Mann Kendall (M.K.) trend analysis, colored boxes show a significant trend (b).

3.2. Model output evaluation

3.2.1. Soil moisture

To evaluate the ability of the ILDAS to simulate soil moisture, we calculated the coefficient of correlation and unbiased RMSE (Fig. 4) of simulated monthly mean soil moisture anomalies with ESA-CCI monthly mean soil moisture anomalies for the period of 2007 to 2021. The period is selected based on continuous data availability without gaps over India. After evaluating the basin-wise coefficient of correlation median values over the primary basins, we found that MERRA-2 shows a high correlation in most of the basins, with the highest in the west flowing rivers from Tapi to Kanyakumari (0.95) (Fig. 4a). Kantha Rao and Rakesh, (2019) also found a high correlation of simulated soil moisture and in-situ observations in the coastal regions. We found that MERRA-2 and ERA-5 show less ubRMSE compared to CHIRPS and IMD in most of the basins (Fig. 4b).

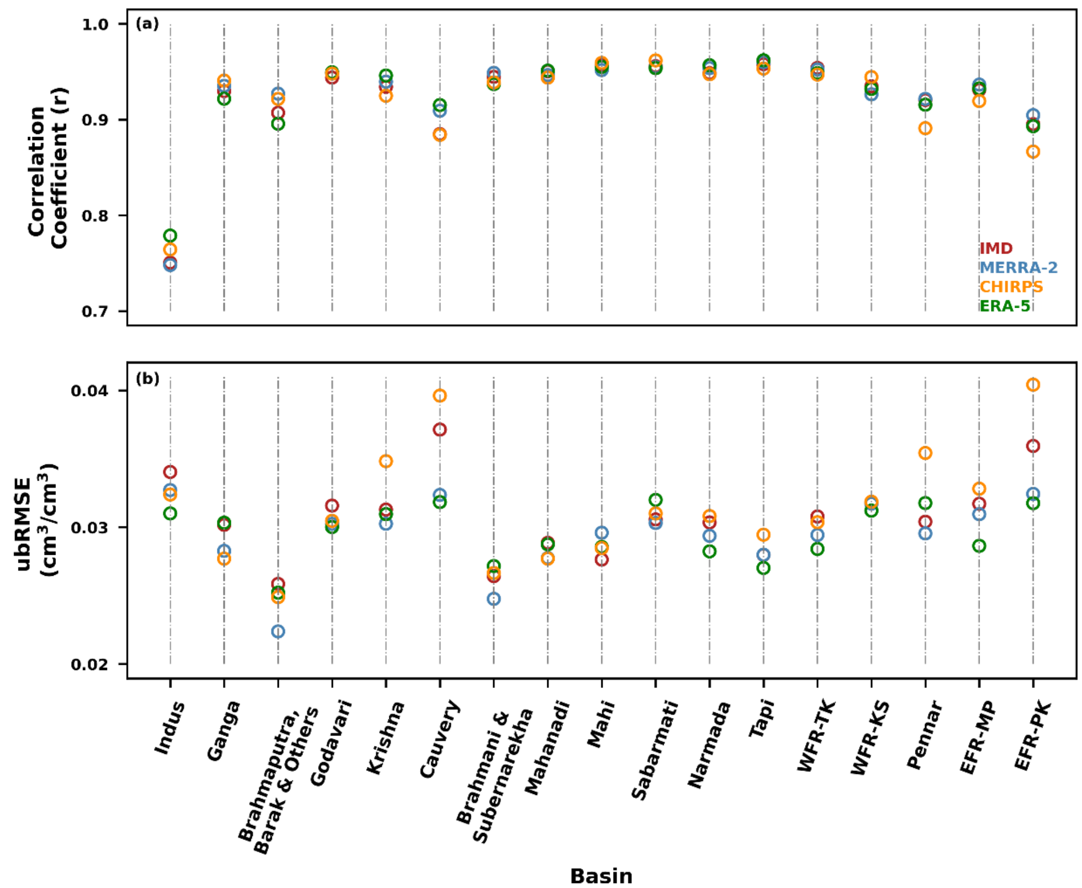


Fig. 4. Basin-wise comparison of correlation, and unbiased RMSE of simulated monthly soil moisture anomalies for different meteorological forcings (IMD, MERRA-2, CHIRPS, and ERA-5) with ESA-CCI soil moisture anomalies for 2007-2021.

The in-situ soil moisture observations in India are rare due to sparse gauge network and limited data availability. The validation with in-situ COSMOS soil moisture observations was performed at daily and monthly scale for a period of 2015-2019 at two-gauge stations. The monthly simulated soil moisture shows good agreement with in-situ soil moisture data throughout the time series as the R^2 is greater than 0.66 at both gauge stations for all four meteorological forcing (Fig. 5). The model retains skill at daily scale as R^2 varies from 0.74 to 0.68 and 0.61 to 0.57 at SGR and MDH, respectively (Fig. 6). We also evaluated basin-wise trend and found that the simulated and ESA-CCI soil moisture anomaly do not show significant trend in most of the basins at monthly scale (Fig. S1).

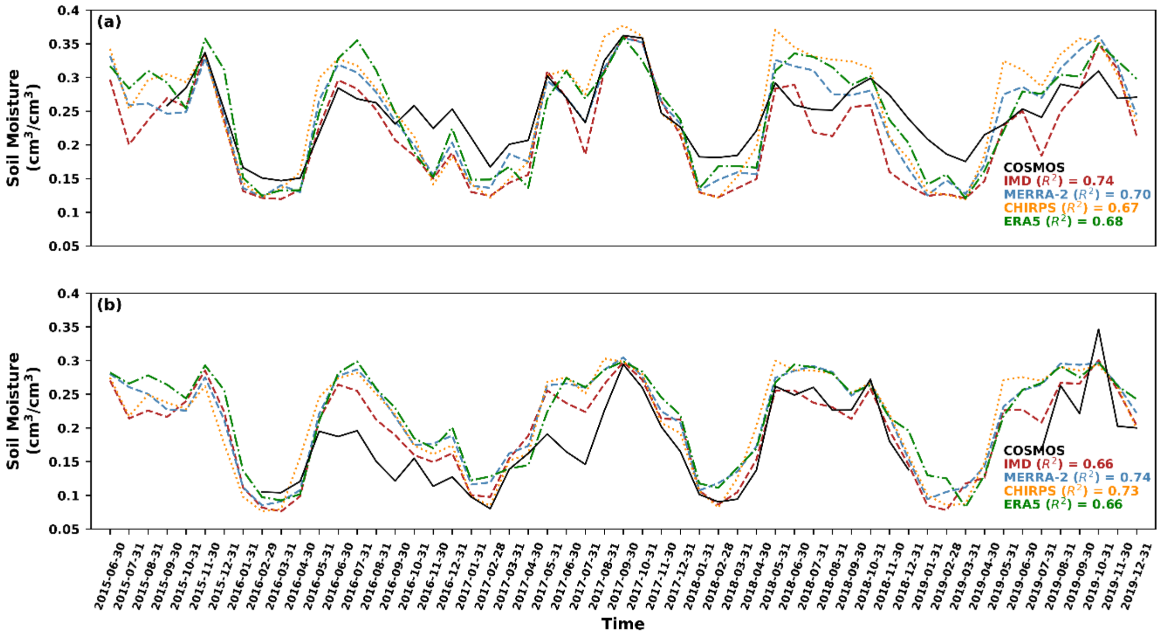


Fig. 5. Comparison of simulated monthly mean soil moisture for different meteorological forcings (IMD, MERRA-2, CHIRPS, and ERA-5) with COSMOS (in-situ) soil moisture at two-gauge stations (a) SGR and (b) MDH from June 2015 to December 2019.

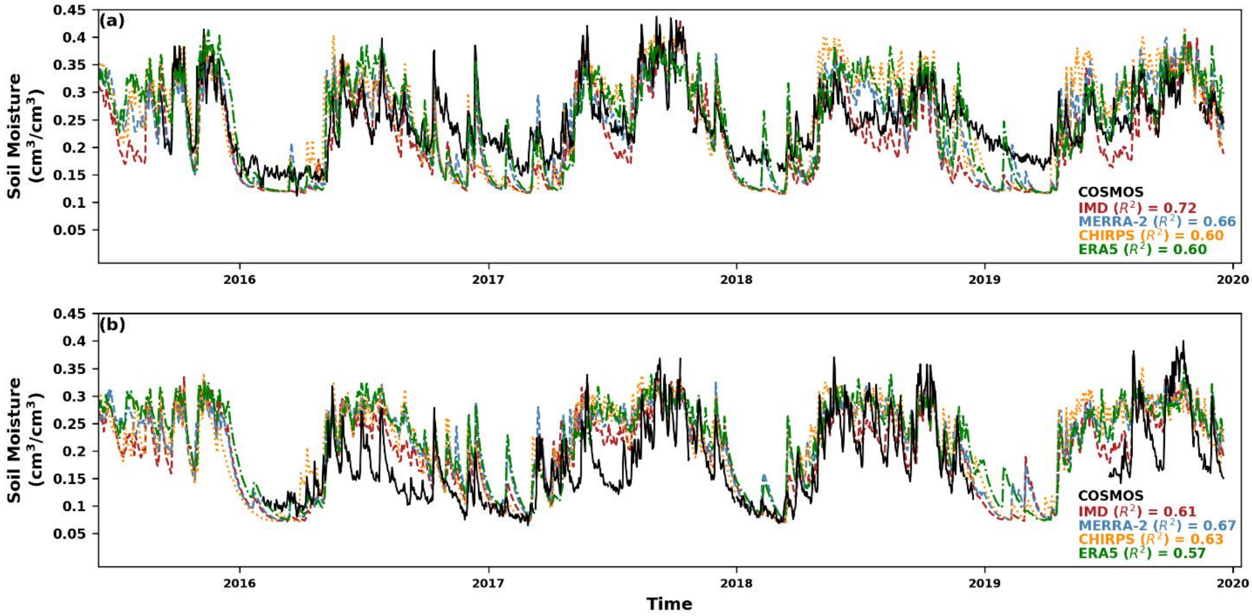


Fig. 6. Comparison of simulated daily mean soil moisture for different meteorological forcings (IMD, MERRA-2, CHIRPS, and ERA-5) with COSMOS (in-situ) soil moisture at two-gauge stations (a) SGR and (b) MDH from June 2015 to December 2019.

3.2.2. Evapotranspiration

We evaluated the ILDAS simulated evapotranspiration with the four meteorological forcings against the MODIS Aqua Evapotranspiration from 2002 to 2021 (Fig. 7). We calculated basin-wise KGE (Fig. 7a) and found that the ILDAS performs well in most of the basins except Indus, Cauvery, and east flowing rivers from Pennar to Kanyakumari. ERA-5 shows a high basin-wise median r (Fig. 7b) in most of the basins than IMD, MERRA-2 and CHIRPS. Most of the basins show variability greater than one for evapotranspiration with all four meteorological forcings (Fig. 7c). However, Brahmaputra and west flowing rivers from Tapi to Kanyakumari show variability near to one than other basins (Fig. 7c). We found that all four meteorological forcings (IMD, MERRA-2, CHIRPS, and ERA-5) show overestimation in most of the basins except Brahmaputra, and west flowing rivers from Tapi to Kanyakumari (Fig. 7d). However, in an early study, Srivastava et al., (2017) stated that the

MODIS underestimated evapotranspiration in India; Our results show a higher bias in most of the basins. All four meteorological forcings show overestimation of the mean annual evapotranspiration compared to the MODIS aqua evapotranspiration (Fig. S2a). Most of the basins show a positive trend in simulated and MODIS aqua evapotranspiration except Brahmaputra, whereas CHIRPS shows a negative trend (Fig. S2b).

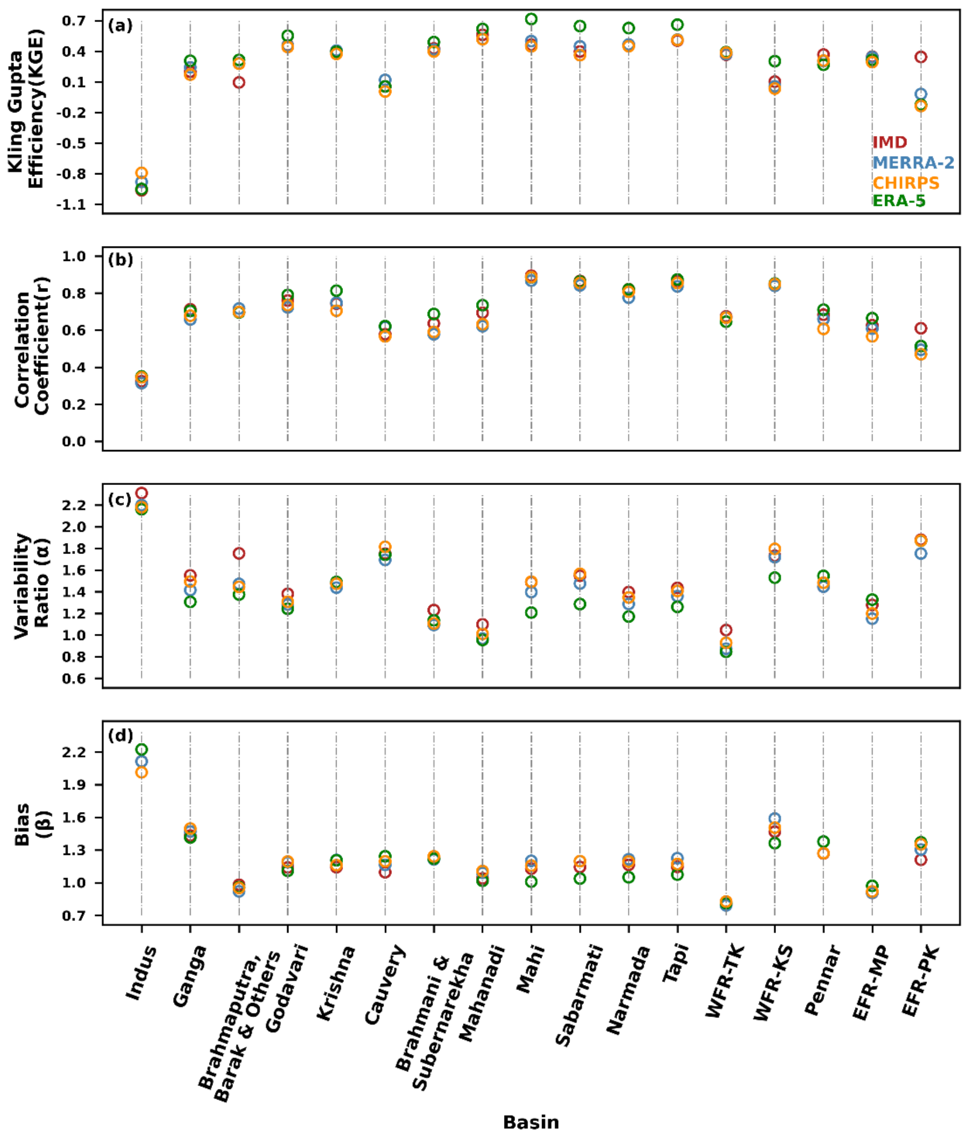


Fig. 7. Basin-wise comparison of KGE, correlation, α, and β of simulated monthly mean evapotranspiration for different meteorological forcing (IMD, *MERRA-2, CHIRPS, and ERA-5) with MODIS Aqua (MYD16A2GF) Evapotranspiration product for 2002-2021.*

3.2.3. Runoff

We evaluated simulated runoff for different meteorological forcings against GRUN runoff (Fig. 8a-d). We found that all four meteorological forcings show poor performance (KGE<0.2) in the Indus and Brahmaputra River basins (Himalayan region) (Fig. 8a). However, the ERA-5 performs better other three forcings in Indus and Brahmaputra. GRUN (Ghiggi et al., 2019) did not consider glacier melting in the generation of runoff data and this may lead to larger uncertainties in the Himalayan region, and our meteorological forcings underestimated the precipitation in this region, which may incorporate the uncertainty in the runoff. All four meteorological forcings show a high correlation in the Indian subcontinent (Fig. 8b). However, the correlation is relatively less in the Indus and Himalayan regions compared to other parts of India. Next, we checked the basin-wise variability of simulated runoff from the model with all four meteorological forcings against the GRUN runoff (Fig. 8c). We found that variability in the runoff is less than one in the basins for all four meteorological forcings except Tapi. However, ERA-5 and IMD show variability closer to one compared to MERRA-2 and CHIRPS. We found that the runoff is highly underestimated (Fig. 8d) in the Indus, Brahmaputra, and Western Ghats regions, possibly due to the hilly terrains in these regions and the lack of incorporation of irrigation practices in our current system. Irrigation leads to a decrease in runoff, which is currently underrepresented and will be incorporated in the next version of the system. We found that all meteorological forcings are underestimating the runoff (Fig. S3a). Next, we evaluated the basin-wise trend, all meteorological forcings showed positive trend in most of the basins (Fig. S3b). Overall, IMD performed better compared to MERRA-2, CHIRPS and ERA-5 in simulating runoff.

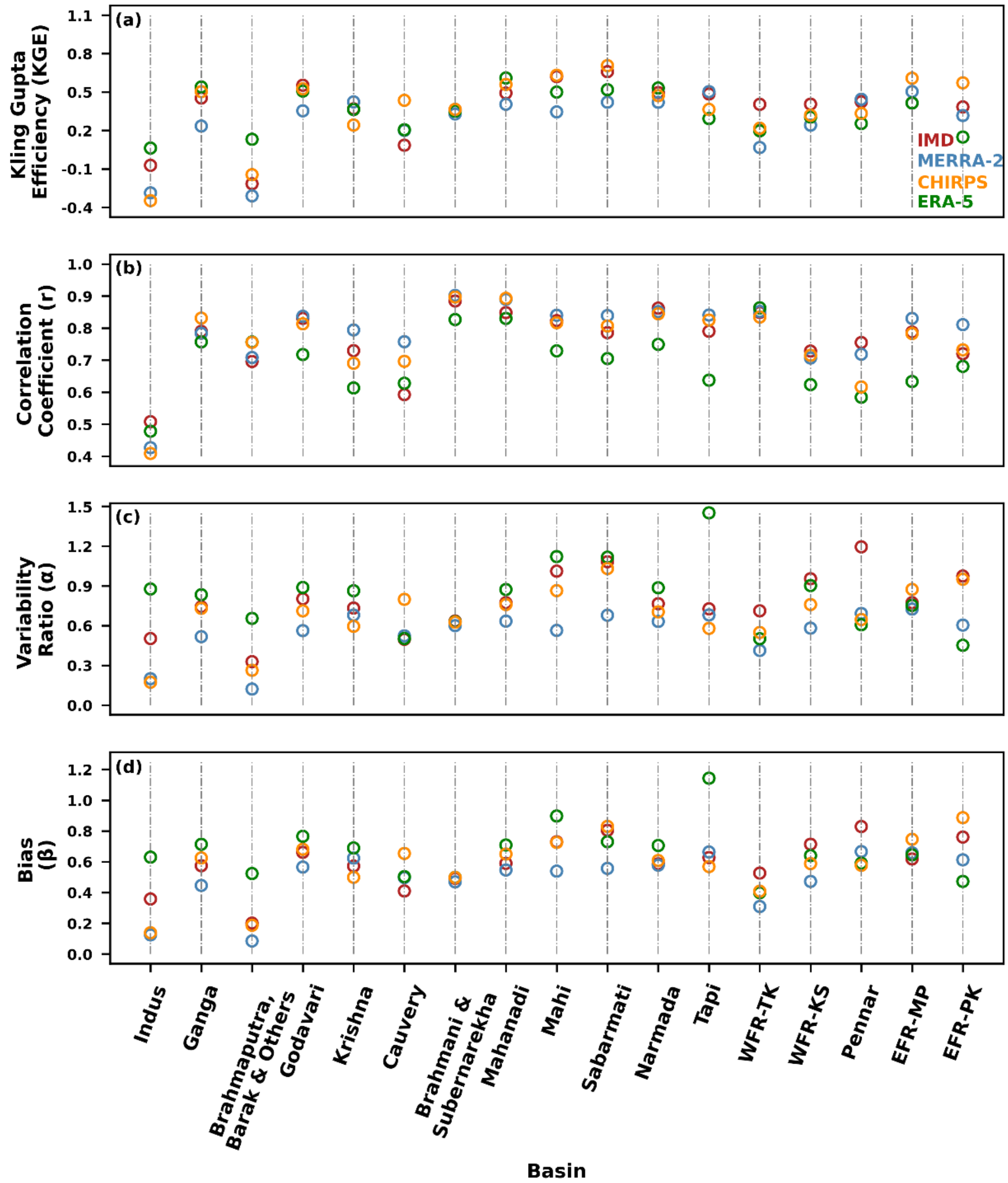


Fig. 8. Basin-wise comparison of KGE, correlation, α, and β of simulated monthly mean Runoff for different meteorological forcings (IMD, MERRA-2, CHIRPS, and ERA-5) with monthly mean GRUN Runoff for 1981-2014.

3.2.4. Streamflow

We calculated KGE and its components for simulated vs observed monthly streamflow from 1981-2021 for each gauge location. The spatial distribution of performance for all gauge stations on an annual basis is shown in (Fig. 9). Comparing the nationwide median and interquartile range (IQR) of overall KGE score, IMD scored the highest median value of 0.36 (IQR: 0.08 – 0.57), closely followed by CHIRPS and ERA-5 with median values of 0.33 (IQR: 0.04 – 0.56) and 0.3 (IQR: -0.08 – 0.58), respectively. MERRA-2 scored the lowest median value of 0.27 (IQR: 0.06 – 0.47). The west flowing rivers from Tapi to Kanyakumari show the highest KGE scores, whereas the gauge stations in central India exhibited majority of the underperformance (Fig. 9a-d). While comparing the performance of individual KGE components, we found that all four forcings showed a good median correlation ($r > 0.7$), with IMD scoring the highest nationwide median value of 0.83, followed by ERA-5 (0.81), CHIRPS (0.75) and MERRA-2 (0.71). Additionally, 94% of gauge stations had a $r >= 0.5$ for ERA-5, 92% for IMD, 89% for CHIRPS, and 79% for MERRA-2. It may be noted that even though the median correlation for ERA-5 is lower than IMD, it shows correlation greater than 0.5 in more basins compared to IMD. The spatial distribution of gauge stations with high r scores matches with those that had high overall KGE scores, with most of the underperformance seen in upper Ganga River basin (Fig. 9e-h) as multiple reservoirs and other irrigation structures result in a delayed response in the river's streamflow. In terms of statistical variability of monthly flows, IMD, CHIRPS and ERA-5 had $\alpha > 1$ in most gauge stations (51%, 52% and 61%, respectively), which corresponds to higher variability in simulated values as compared to the observed ones. In contrast, MERRA-2 showed low variability with $\alpha < 1$ in 63% of the gauge stations. In terms of the spatial distribution, the relatively lower values of α are seen majorly in Ganga River basin and some of the gauge stations in west flowing rivers. However, for ERA-5 simulations, the variability in Ganga River basin is higher than the observed, which also resulted in the highest overall median α (1.08). The nationwide median β was lowest for MERRA-2 (1.11), followed by IMD (1.13), CHIRPS (1.15) and ERA-5 (1.4). All forcings showed positive bias in simulated streamflow for most of the stations with ERA-5 showing highest number of gauge stations with overestimated streamflow (70%), followed by CHIRPS (61%), IMD

(59%) and MERRA-2 (53%). Additionally, we also found that ERA-5 had the highest number of gauge stations (20%) with a bias greater than 100% (β >=2), compared to CHIRPS (12%), IMD (9%) and MERRA-2 (4%). The high median β for streamflow simulated by ERA-5 and CHIRPS shows that the ILDAS struggled to match the magnitude of seasonal flows, especially in the central and peninsular regions which is expected due to the non-perennial rivers and various anthropogenic activities (Fig. 9m-p). Overall, IMD can be considered as the best performing precipitation forcing among the four based on median KGE value.

Besides annual evaluation, since most of the precipitation over India is concentrated in the months of June-September (also known as JJAS season) which results in very high seasonal flows in the Indian rivers, we also evaluated the simulated streamflow specifically for JJAS season (Fig. 10). Overall, the nationwide KGE median score increased by 11.1% for MERRA-2 and IMD) and 6.6% for ERA-5 (0.32 vs 0.3) but decreased for CHIRPS by 6% (0.33 vs 0.31). However, all four forcings saw a reduction of in median r score in JJAS season, with values of 0.62, 0.64, 0.71, 0.77 for MERRA-2, CHIRPS, ERA-5 and IMD, respectively. Additionally, a corresponding reduction in β is observed, while α decreased marginally for CHIRPS but increased for MERRA-2, ERA-5 and IMD. Hence, during the JJAS season, ILDAS captured the magnitude and variability of high monsoon flows with a higher skill, but the timing could not be matched well, which is due to the various regulatory structures such as reservoirs resulting in a reduced as well as delayed streamflow in the rivers.

The performance of the integrated hydrological-hydrodynamic model can be assessed by evaluating the overall patterns of streamflow. Therefore, we calculated monthly streamflow anomalies for all four forcings and compared them against the observed to assess the ability of ILDAS to capture general streamflow patterns during the annual cycle and monsoon season. We used the anomaly correlation coefficient and unbiased RMSE (ubRMSE) to evaluate the performance of the model across 162 catchments (Fig. 11-12). The results of the annual evaluation showed that the IMD driven streamflow had the highest correlation with the observed anomalies (0.69), followed by ERA-5 (0.57), MERRA-2 (0.53), and CHIRPS (0.51). Although the anomaly correlation coefficient marginally improved during the JJAS season, the ubRMSE showed a significant increase in value across all forcings, suggesting that ILDAS overestimated the anomalies during the monsoon (Fig. 12e-h). This could be due to the lack of information regarding various management practices, such as reservoirs, in ILDAS which caused the model to simulate higher flows than observed. The overall superior performance of IMD could be due the localized and more accurate precipitation information as it leverages the extensive network of rain gauges across India.

On the daily scale, daily streamflow shows a nationwide median KGE of 0.27, compared to 0.36 for IMD monthly. The error metrics for other meteorological forcings are presented in Figure S5-S6. The assessment of daily streamflow anomaly correlation for annual season shows that IMD has highest correlation with observed anomalies (R = 0.48), followed by ERA-5 (0.36), CHIRPS (0.31), and MERRA2 (0.29). For JJAS season, the daily anomaly correlation largely remains same but ubRMSE increases significantly, indicating higher variability in daily monsoon flows (Fig.13). The skill of streamflow simulations at daily scale emphasizes the future need for calibration and including anthropogenic effects into the model such as reservoirs.To further investigate the performance of ILDAS, we also calculated commonly used hydrological signatures such as mean annual flow, mean annual monsoon flow, low flow, and high flow. Using the coefficient of determination (R^2) as the performance metric, we observed that ERA-5 had the highest R^2 scores across all hydrological signatures followed by CHIRPS, IMD, and MERRA-2 (Fig. 14). The highest and lowest R^2 scores were observed for mean annual high flow and mean annual low flow, respectively. The low R^2 for mean annual low flows emphasizes the need for incorporating the anthropogenic effects and calibration of the model.

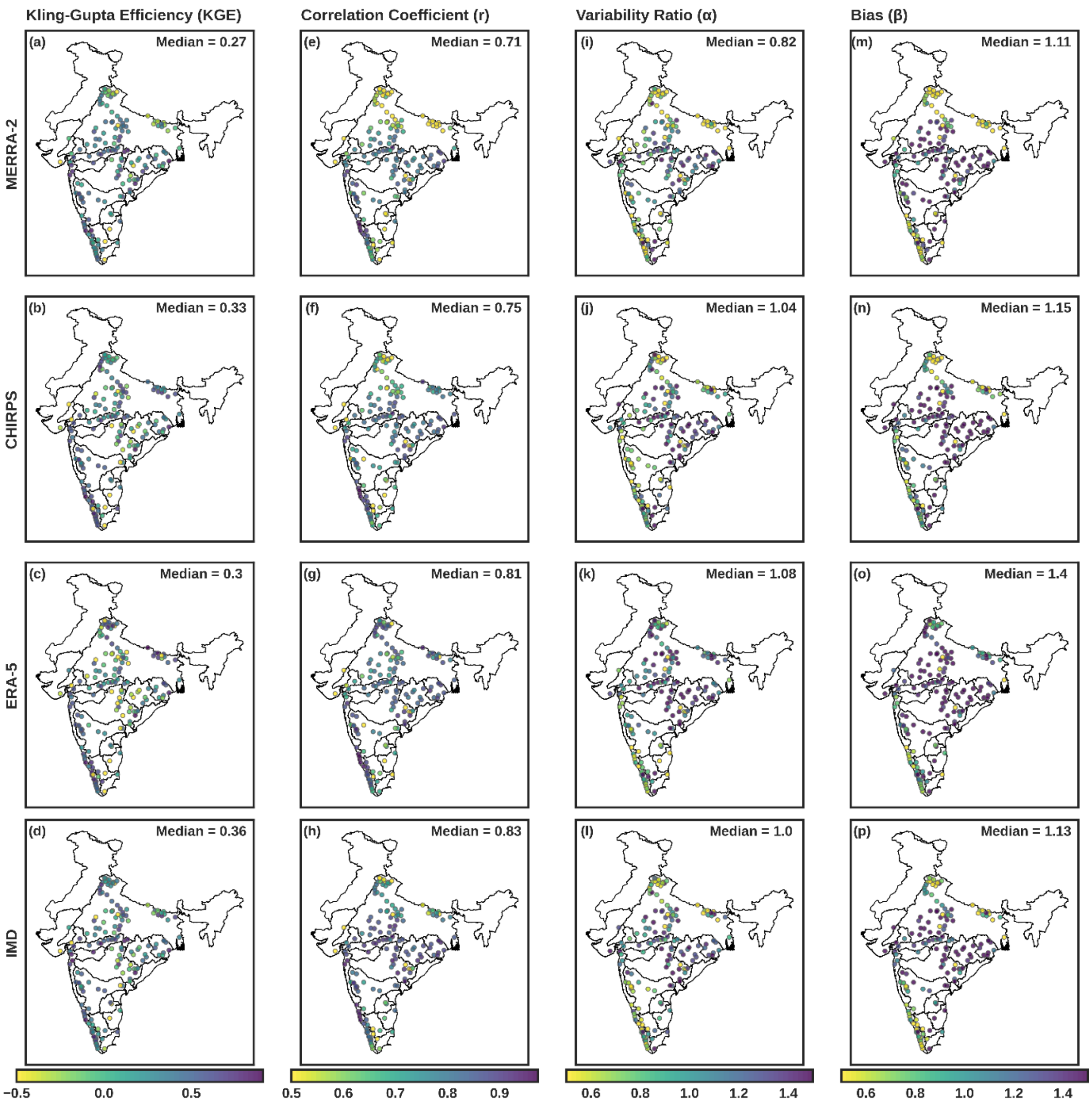


Fig. 9. Comparison of KGE (a-d), r (e-h), α (i-l), and β (m-p) of simulated monthly mean streamflow annually for different meteorological forcings at 162 gauge stations.

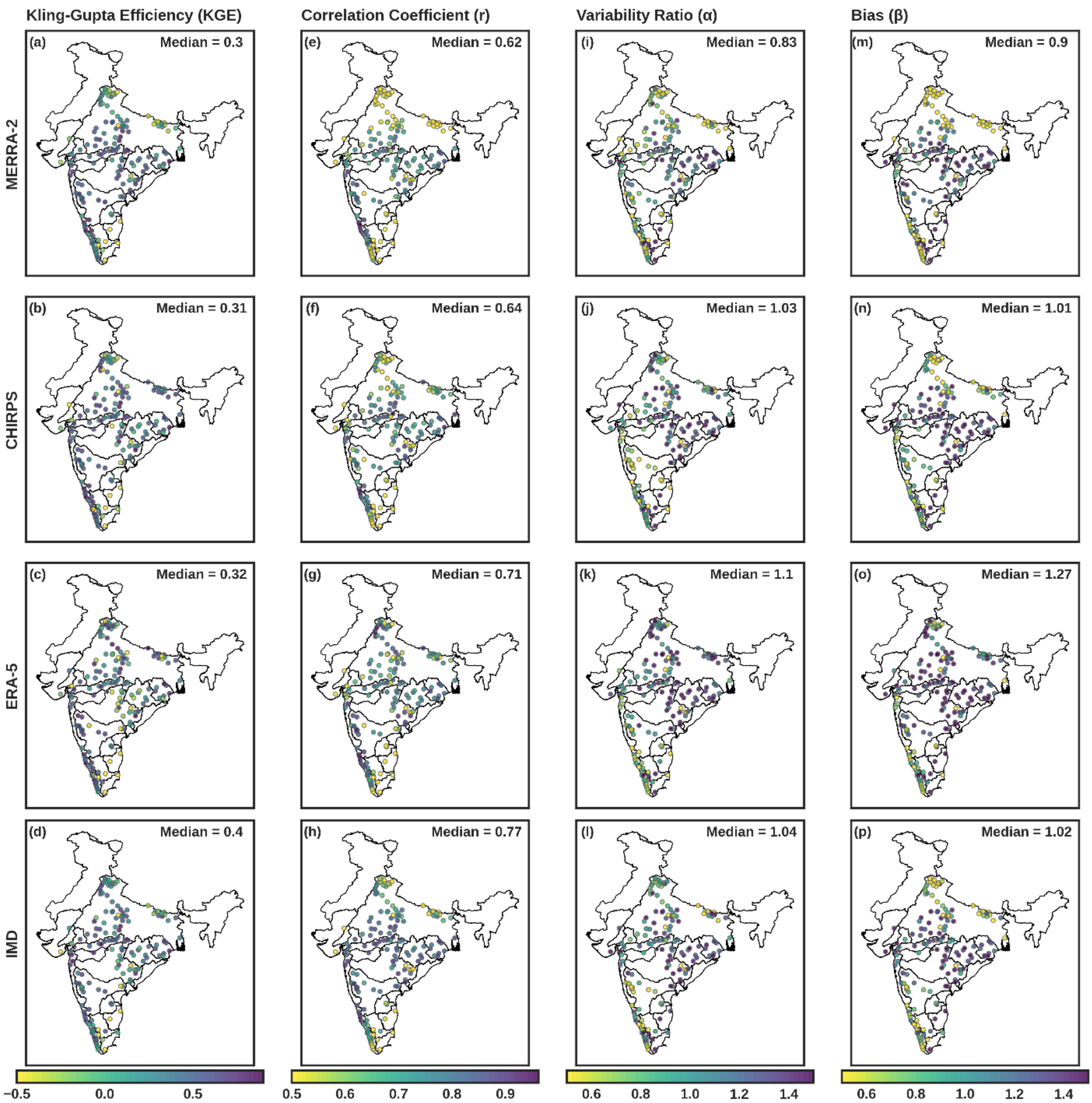


Fig. 10. Comparison of KGE (a-d), r (e-h), α (i-l), and β (m-p) of simulated monthly mean streamflow in JJAS season for different meteorological forcings at 162 gauge stations.

Table 5. Detailed analysis of monthly simulated streamflow against observed streamflow for annual and JJAS (monsoon season) from 1981-2019 for the four forcings. The digits represent the number of gauge stations (out of 162) falling under the specified criteria.

Criteria	**MERRA-2**	**CHIRPS**	**ERA-5**	**IMD**

	Annual	JJAS	Annual	JJAS	Annual	JJAS	Annual	JJAS
KGE Inter Quartile Range	0.06 - 0.47	0.01 - 0.51	0.04 - 0.56	0.11 - 0.49	-0.08 - 0.58	0.02 - 0.52	0.08 - 0.57	0.1 - 0.59
Correlation (r) Distribution (>=0.5/<0.5)	128/34	113/49	144/18	129/33	152/10	142/20	147/15	135/27
Variability (α) Distribution (low/high)	103/59	101/61	78/84	78/84	64/98	71/91	80/82	73/89
Bias (β) Distribution (negative/positive)	76/86	86/76	63/99	80/82	49/113	54/108	57/95	78/84
β >= 2	7	6	19	9	32	16	14	7

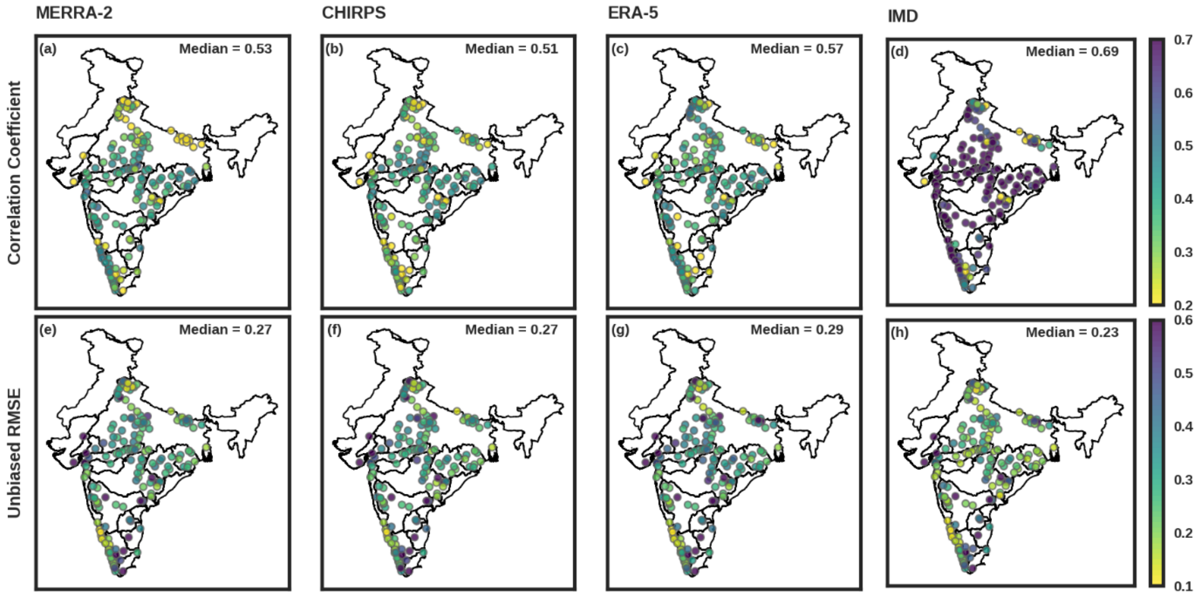


Fig. 11. Anomaly correlation coefficient (a-d) and unbiased RMSE (e-h) for monthly mean streamflow in annual season for different meteorological forcings at 162 gauge stations.

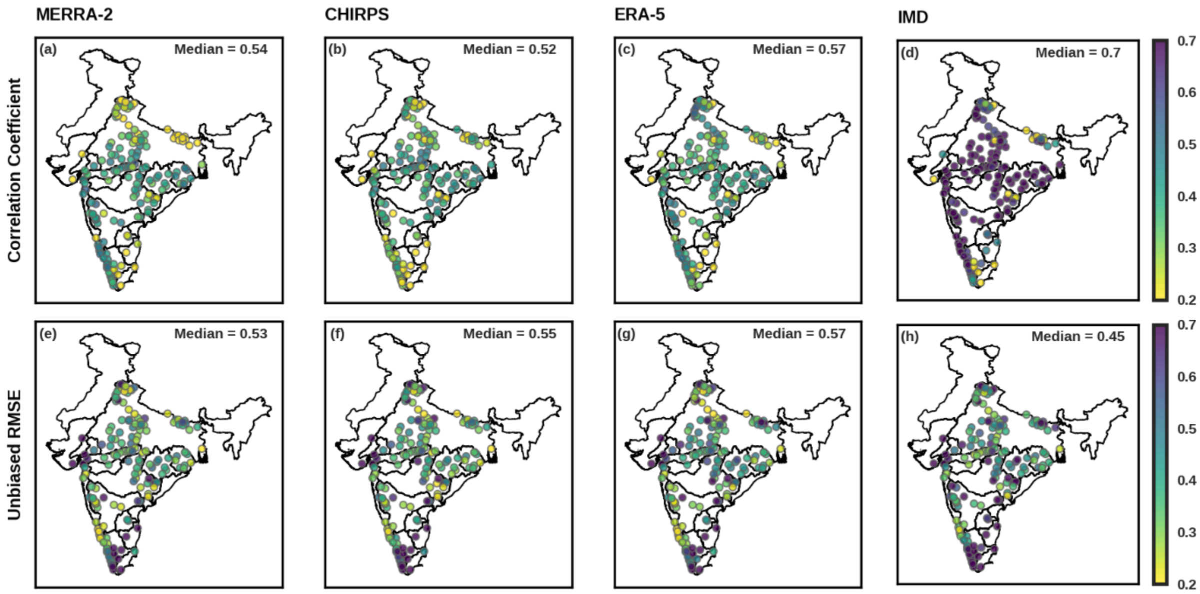


Fig. 12. Anomaly correlation coefficient (a-d) and unbiased RMSE (e-h) for monthly mean streamflow in JJAS season for different meteorological forcings at 162 gauge stations.

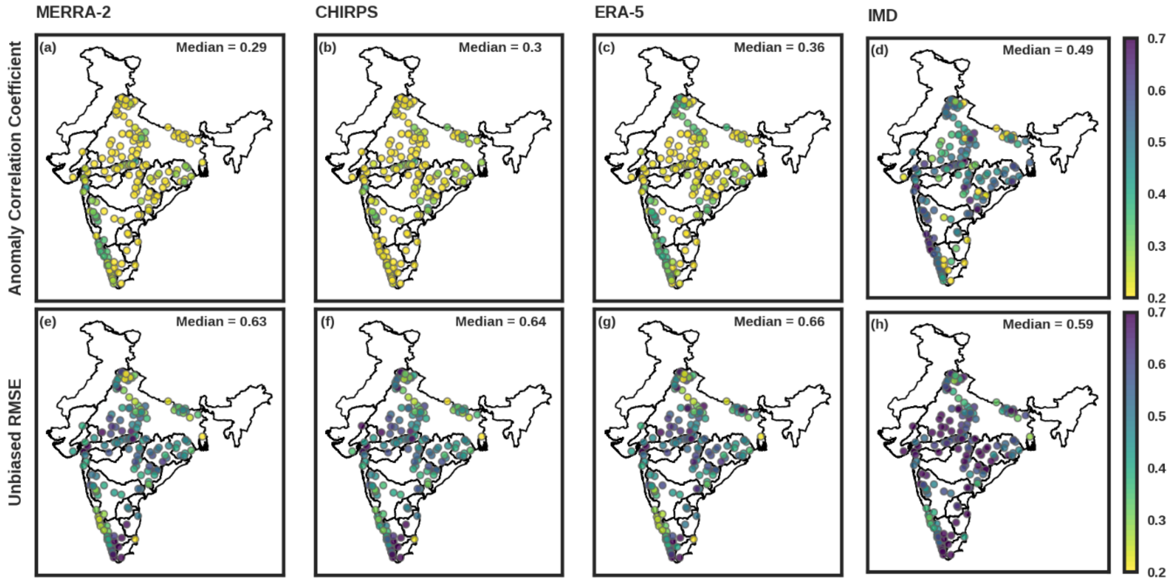


Fig. 13. Anomaly correlation coefficient (a-d) and unbiased RMSE (e-h) for daily streamflow in JJAS season for different meteorological forcings at 162 gauge stations.

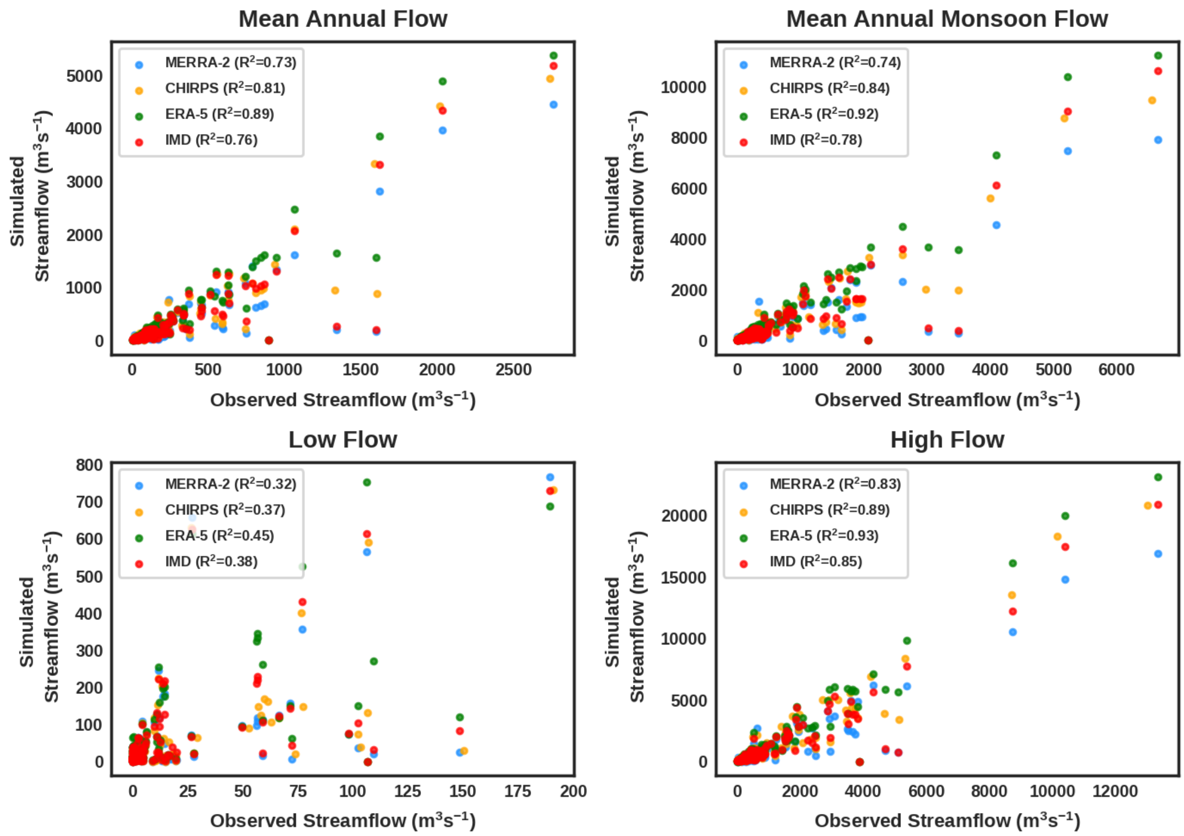


Fig. 14. Scatter plots of various hydrological signatures for 162 catchments across Indian subcontinent. The values in-set denote coefficient of determination (R^2) corresponding to each hydrological signature.

3.2.5. Terrestrial Water Storage Anomaly (TWSA)

We evaluated the TWS for all meteorological forcings by adding Land Surface Model (LSM) water storage (LWS) and Surface water storage (SWS). LWS consists of groundwater storage (GWS), soil moisture (SM), and snow water equivalent (SWE). Most of the studies do not consider the SWS in the TWS. However, Surface water storage (SWS) contributes 8% of TWS variability globally (A. Getirana, Kumar, et al., 2017). While comparing the time series of nationwide monthly mean simulated TWSA using Noah MP + HyMAP with meteorological forcings to GRACE TWSA for 2003-2017 (Fig. 15a), we found that all four meteorological forcings had captured the seasonality of TWSA well. However, from 2010 onward, all meteorological forcings overestimated the peaks and the troughs. We noted that GRACE shows a negative trend in the TWSA. This negative trend may be due to the extensive extraction of groundwater in parts of India such as Punjab, East Flowing River (Pennar-Kanyakumari) (EFR-PK) and Ganga basin. Similar patterns were observed by previous studies in the Indian mainland (Satish Kumar et al., 2023). We found that CHIRPS shows a positive trend, whereas IMD, MERRA-2, and ERA-5 show relatively no trend. Uncertainties in the TWSA GRACE and ILDAS simulated TWSA may be due to India's anthropogenic conditions and irrigation, which will be incorporated in ILDAS in the future. We also calculated the basin-wise R^2 for the primary basin and found that most of the basins show high R^2 for MERRA-2 and ERA-5 (Fig. 15b). Moreover, all forcings show poor R^2 in the Indus River basin which could be due to the excessive groundwater extraction in this region. Previous studies (Asoka & Mishra, 2020; Maina et al., 2022) have also shown a similar pattern in the northwest (Indus) region. Overall, our results show that IMD, MERRA-2, and ERA-5 performed well with a nationwide mean (R^2 >0.57) except CHIRPS (R^2 = 0.53). Similarly, Soni and Syed, (2015) also found similar performance in the major river basins of India.

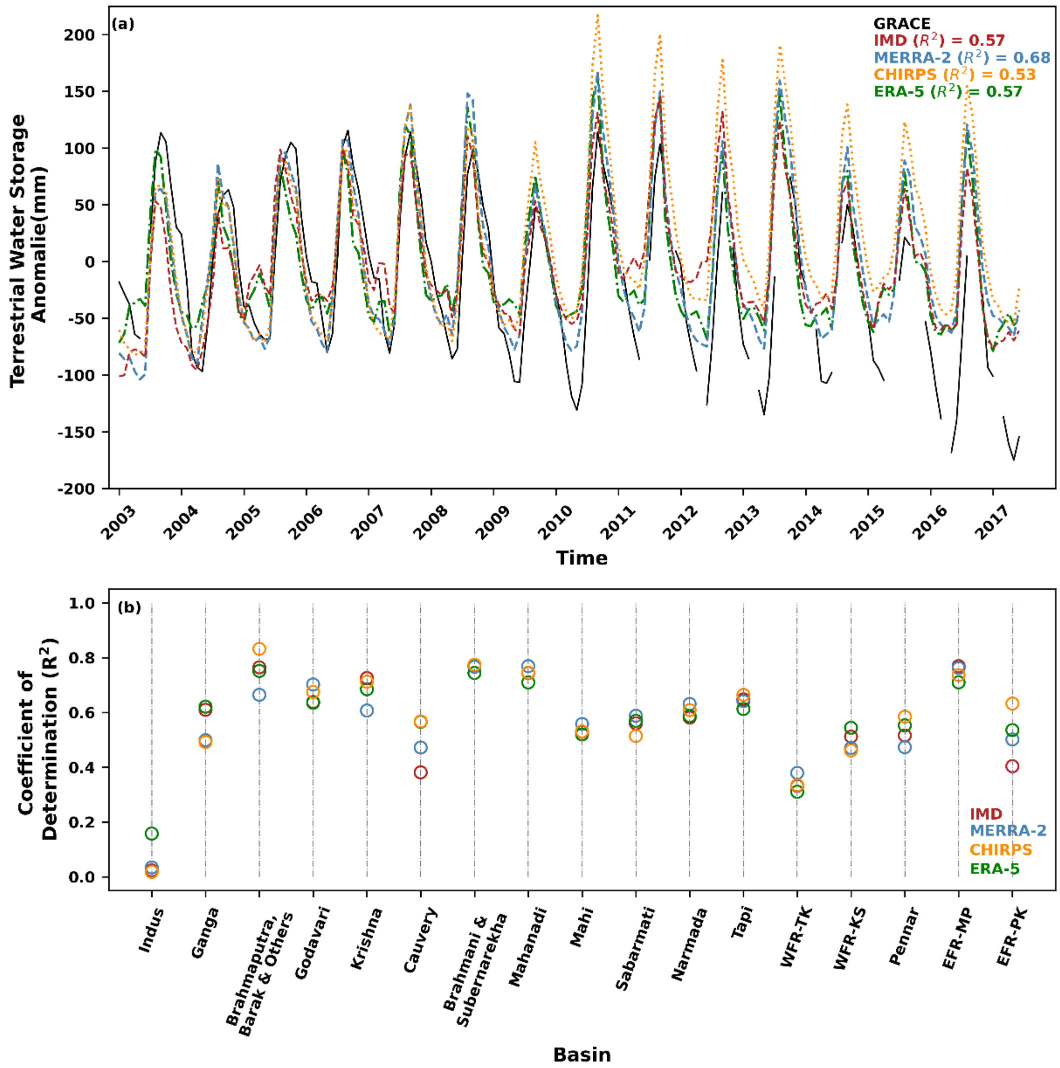


Fig. 15. A plot of time series plot for nationwide monthly mean terrestrial water storage anomaly for GRACE, IMD, MERRA-2, CHIRPS, and ERA-5 for 2003-2017 (a) and basin-wise R^2 (b).

3.2.6. Seasonal Water Balance Cycle

A coupled hydrological-hydrodynamic model is expected to capture the variation of long-term water balance of the region. Therefore, along with the quantitative assessment of water balance components discussed in the previous sections, we've tried to illustrate the ability of ILDAS in capturing the seasonal variation of the terrestrial water budget using time series plots of various water balance components along with the anomalies of water fluxes and the terrestrial water storage. Here, we present the qualitative analysis for simulated water balance using CHIRPS at Kudige, Cauvery River basin, which is a rain-fed region in the southern India (Fig. 16-17). Fig. 16 shows the long-term variation of various water balance components for the period 1981-2021. Additionally, Fig. 17 shows the monthly anomalies for simulated water balance for four different meteorological seasons. We observed that ILDAS is successful in simulating the long-term seasonal variation in terrestrial water storage with precipitation as the primary factor. The terrestrial water storage remains in deficit compared to long term monthly mean when precipitation is low in winters and summers, followed by a surplus period in monsoons, which agrees with the climate and topography of the basin. Moreover, the deficit is largest during the peak summer which gets replenished in the subsequent monsoon.

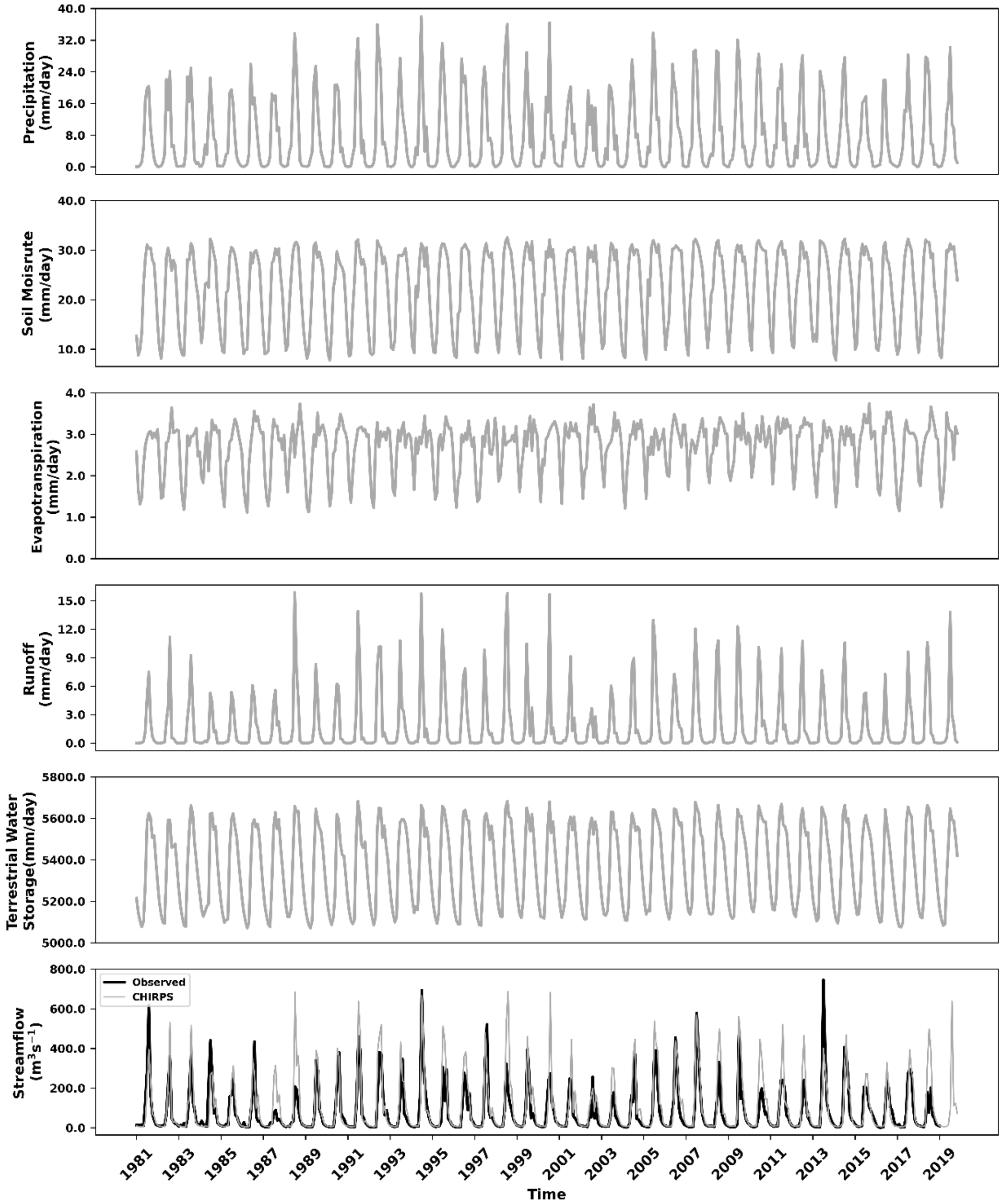


Fig. 16. A time-series plot showing long-term variation of various components of simulated water balance at Kudige, Cauvery River basin.

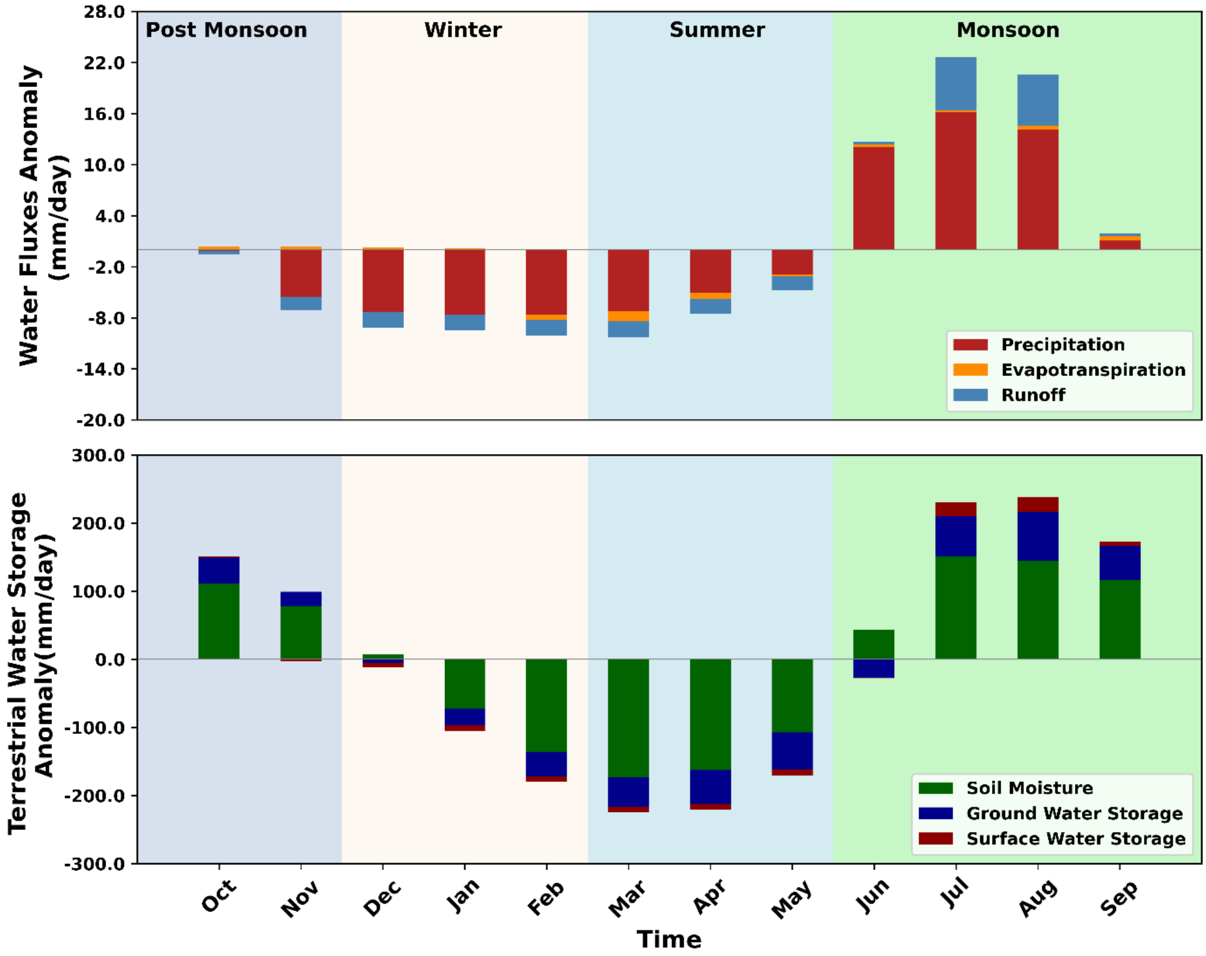


Fig. 17. Seasonal variations of different water balance components as monthly anomalies simulated using ILDAS forced by CHIRPS at Kudige, Cauvery River basin.

3.2.8 Hydroclimatic Extreme Event Analysis

In May 2022, the town of Haflong, located in the Dima Hasao district of Assam, India, experienced a catastrophic series of landslides and floods resulting in extensive damage and loss of life and property. The disaster occurred between May 11-18 and was triggered by heavy rainfall, affecting multiple villages in the area (Roy et al., 2023). The landslides caused severe damage to infrastructure such as roads, bridges, and buildings, hampering rescue and relief efforts. To better understand the underlying conditions, we reconstructed the total column soil moisture (1000 mm) in the area from 1981-2022 and compared the 2022 daily soil moisture anomaly to 1981-2021 median (Fig. 18). Fig. 18

illustrates that the antecedent soil moisture anomaly in the area was significantly higher than the long-term median, indicating saturation of the soil due to heavy rainfall on April 15-17. This heightened soil moisture content increased the vulnerability to landslides and inundation. The subsequent high rainfall in May resulted in high runoffs and increased pore pressure which caused district wide inundation and cluster landslides due to slope failure at multiple sites. This finding underscores the importance of monitoring soil moisture conditions and incorporating this information into landslide risk assessment and management strategies. The ILDAS was able to capture the local antecedent soil moisture condition even at an uncalibrated stage which is a promising prospect for future implementation in operational forecasts.

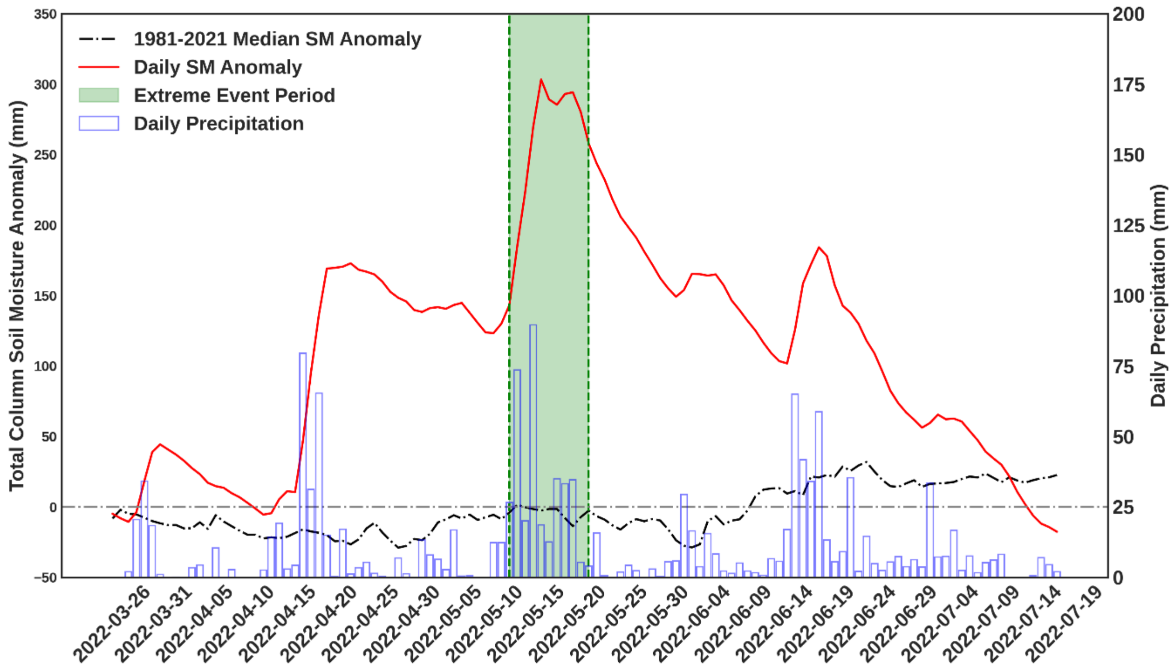


Fig. 18. Reconstructed daily soil moisture anomaly in Haflong, Assam for the extreme hydroclimatic event that occurred from 11-18 May 2022.

4. Conclusions

We have established ILDAS as a prototype of a coupled hydrologic-hydrodynamic system to generate a high-quality reanalysis of land surface estimates and streamflow at 0.1° resolution and daily temporal resolution across the Indian mainland for the period 1981-2021. We tested the ILDAS using three meteorological forcings with varying spatial and temporal resolutions and assessed its ability to simulate various water balance components such as soil moisture, evapotranspiration, surface runoff, streamflow, and terrestrial water storage anomalies. We evaluated the uncertainty and bias in the precipitation component of three global meteorological forcings across the various regions of India. We found that CHIRPS exhibits lower uncertainty than MERRA-2 and ERA-5, and a high correlation and minimum RRMSE against observation-based IMD precipitation. Additionally, we evaluated all major components of simulated water balance. It was found that all meteorological forcings showed good performance for simulated soil moisture by ILDAS. However, MERRA-2 showed minimum median ubRMSE value for most of the basins compared to others. The correlation is high in simulated soil moisture as well as runoff for all meteorological forcings. However, our results did not indicate good agreement with GRUN runoff in the Himalayan region, which may be due to glacier melting not being considered in generating GRUN runoff. We evaluated the average monthly streamflow against observed streamflow at multiple gauge stations for annual and the monsoon (JJAS) season. The overall results from the annual evaluation show that ILDAS could match the streamflow timing better than the magnitude and variability, especially in central and peninsular India, which may be caused by very low flows during the non-monsoon months in seasonal rivers, that are further reduced due to human abstractions. However, while evaluating the streamflow specifically during the monsoon months, we found that the overall nationwide median values of both r and β were reduced. This means that although ILDAS performance improved in capturing the magnitude of streamflow, the timings of the flows could not be matched well. However, an overall nationwide KGE of 0.36 for IMD in annual evaluation is a promising result for ILDAS, which can be further improved using data assimilation, calibration, or post-processing. The evaluation of monthly streamflow anomalies for annual season agreed with the evaluation of streamflow timeseries and IMD showed highest anomaly correlation coefficient and lowest ubRMSE. However, ubRMSE degraded for all the forcings in JJAS season due to the uncalibrated state of ILDAS, resulting in higher than observed flows. While evaluating the streamflow-derived hydrological signatures, we observed that IMD's superior performance in simulating temporal patterns of streamflow did not translate to overall statistical streamflow patterns of the catchments. The global forcings including ERA-5 and CHIRPS performed better in simulating the hydrological signatures compared to IMD. The overall evaluation of water balance components suggests that different meteorological forcings performed better for different land surface variables, which highlights the value of developing an ensemble of model configurations with multiple data sources. We also reconstructed antecedent soil moisture during a recently occurred hydro-climatic extreme in Haflong town of Assam, India, causing a series of landslides and inundation in the area. We observed that ILDAS successfully simulated the observed daily soil moisture anomaly, which was significantly higher before the extreme precipitation that occurred in May. This highlights the importance of a high resolution hydrological-hydrodynamic model, such as ILDAS, for risk assessment and disaster mitigation of hydro-climatic extremes.

This study is envisioned as a proof-of-concept of an integrated system over an underserved region such as the Indian subcontinent as most land surface models run in uncoupled mode with river routing models. The ILDAS will serve as a testbed for future experiments on assimilating remote sensing observations and provide near real-time estimates of land surface states, natural water balance, and energy fluxes that are consistent across space and time, with the potential to assist policymakers in decision-making related to food security, water resources management, mitigation of natural hazards, and assessing climate change impacts. Furthermore, there is a pressing need for a transboundary water modeling system which can be used by countries to assess inflows and outflows from the Ganga and Brahmaputra, leading to better cooperation within South Asia in the water sector. The first version of ILDAS has some limitations that will serve as the basis for future improvements. Currently, ILDAS outputs are based on a "natural" terrestrial state, as no information regarding irrigation and reservoirs has been incorporated. Moreover, we acknowledge the limitations of simplified assumptions and inaccuracies in parameterization while representing the physical processes. Future enhancements to ILDAS will include data assimilation of remote sensing products, localized land use/land cover parameters, and representation of reservoirs.

Acknowledgements

This research was conducted in the HydroSense lab (https://hydrosense.iitd.ac.in/) of IIT Delhi and the authors would like to gratefully acknowledge IIT Delhi High Performance Computing facility for providing computational and storage resources. Dr. Manabendra Saharia gratefully acknowledges financial support for this work from ISRO Space Technology Cell (RP04139). He also acknowledges the MoES Monsoon Mission III (RP04574); IC-IMPACTS (RP04558); and the Coalition for Disaster Resilient Infrastructure (CDRI) Fellowship (RP04569) for continual support for the project. The authors thank Prof. Sekhar Muddu (IISc) and Ms. Deepti Upadhyay for giving access to COSMOS soil moisture data. The authors gratefully acknowledge the Central Water Commission (CWC), National Water Informatics Centre, and the Ministry of Jal Shakti (MoJS) for providing the streamflow datasets used in this study. The authors also acknowledge ISRO NRSC for providing access to the 1:50000 LU/LC dataset.

Compliance with Ethical Standards

The authors declare that they have no conflict of interest.

Data Availability Statement.

The datasets used in this study are available from the following sources:

- IMD precipitation: https://www.imdpune.gov.in/Clim_Pred_LRF_New/
- Streamflow: Centra Water Commission, India, and India WRIS, https://indiawris.gov.in/wris/
- MERRA-2: GMAO, NASA Goddard Space Flight Centre, https://disc.gsfc.nasa.gov/datasets?project=MERRA-2
- CHIRPS: Climate Hazards Centre, UC Santa Barbara, https://data.chc.ucsb.edu/products/CHIRPS-2.0/
- ERA-5: ECMWF, https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era5
- GRUN Runoff: Institute for Atmospheric and Climate Science, ETH Zurich, https://figshare.com/articles/dataset/GRUN_Global_Runoff_Reconstruction/9228176
- ESA-CCI Soil Moisture: European Space Agency and Technische Universität Wien (TUW), https://www.esa-soilmoisture-cci.org/data
- MODIS Evapotranspiration: GSFC, NASA, https://modis.gsfc.nasa.gov/data/dataprod/mod16.php
- GRACE and GRACE-FO: Center for Space Research (CSR), The University of Texas, https://podaac.jpl.nasa.gov/GRACE

REFERENCES

Asoka, A., & Mishra, V. (2020). Anthropogenic and Climate Contributions on the Changes in Terrestrial Water Storage in India. *Journal of Geophysical Research: Atmospheres*, *125*(10). https://doi.org/10.1029/2020JD032470

Attada, R., Kumar, P., & Dasari, H. P. (2018). Assessment of Land Surface Models in a High-Resolution Atmospheric Model during Indian Summer Monsoon. *Pure and Applied Geophysics*, *175*(10), 3671–3696. https://doi.org/10.1007/S00024-018-1868-Z/FIGURES/15

Bates, P. D., Horritt, M. S., & Fewtrell, T. J. (2010). A simple inertial formulation of the shallow water equations for efficient two-dimensional flood inundation modelling. *Journal of Hydrology*, *387*(1–2), 33–45. https://doi.org/10.1016/J.JHYDROL.2010.03.027

Bharti, V., & Singh, C. (2015). Evaluation of error in TRMM 3B42V7 precipitation estimates over the himalayan region. *Journal of Geophysical Research*, *120*(24), 12,458-12,473. https://doi.org/10.1002/2015JD023779

Bhattacharyya, S., Sreekesh, S., & King, A. (2022). Characteristics of extreme rainfall in different gridded datasets over India during 1983–2015. *Atmospheric Research*, *267*, 105930. https://doi.org/10.1016/J.ATMOSRES.2021.105930

Carrera, M. L., Bélair, S., & Bilodeau, B. (2015). The Canadian Land Data Assimilation System (CaLDAS): Description and synthetic evaluation study. *Journal of Hydrometeorology*, *16*(3), 1293–1314. https://doi.org/10.1175/JHM-D-14-0089.1

Chakravorty, A., Chahar, B. R., Sharma, O. P., & Dhanya, C. T. (2016). A regional scale performance evaluation of SMOS and ESA-CCI soil moisture products over India with simulated soil moisture from MERRA-Land. *Remote Sensing of Environment*, *186*, 514–527. https://doi.org/10.1016/j.rse.2016.09.011

Chaudhary, S., Dhanya, C. T., & Vinnarasi, R. (2017). Dry and wet spell variability during monsoon in gauge-based gridded daily precipitation datasets over India. *Journal of Hydrology*, *546*, 204–218. https://doi.org/10.1016/j.jhydrol.2017.01.023

De Almeida, G. A. M., Bates, P., Freer, J. E., & Souvignet, M. (2012). *Improving the stability of a simple formulation of the shallow water equations for 2-D flood modeling*. https://doi.org/10.1029/2011WR011570 de Goncalves, L. G. G., Shuttleworth, W. J., Burke, E. J., Houser, P., Toll, D. L., Rodell, M., & Arsenault, K. (2006). Toward a South America Land Data Assimilation System: Aspects of land surface model spin-up using the Simplified Simple Biosphere. *Journal of Geophysical Research Atmospheres*, *111*(17). https://doi.org/10.1029/2005JD006297

Dorigo, W., Wagner, W., Albergel, C., Albrecht, F., Balsamo, G., Brocca, L., Chung, D., Ertl, M., Forkel, M., Gruber, A., Haas, E., Hamer, P. D., Hirschi, M., Ikonen, J., de Jeu, R., Kidd, R., Lahoz, W., Liu, Y. Y., Miralles, D., … Lecomte, P. (2017). ESA CCI Soil Moisture for improved Earth system understanding: State-of-the art and future directions. *Remote Sensing of Environment*, *203*, 185–215. https://doi.org/10.1016/j.rse.2017.07.001

Ek, M. B., Mitchell, K. E., Lin, Y., Rogers, E., Grunmann, P., Koren, V., Gayno, G., & Tarpley, J. D. (2003). Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model. *Journal of Geophysical Research: Atmospheres*, *108*(D22). https://doi.org/10.1029/2002JD003296

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., & Michaelsen, J. (2015). The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. *Scientific Data*, *2*(1), 150066. https://doi.org/10.1038/sdata.2015.66

Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., Randles, C. A., Darmenov, A., Bosilovich, M. G., Reichle, R., Wargan, K., Coy, L., Cullather, R., Draper, C., Akella, S., Buchard, V., Conaty, A., da Silva, A. M., Gu, W., … Zhao, B. (2017). The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). *Journal of Climate*, *30*(14), 5419–5454. https://doi.org/10.1175/JCLI-D-16-0758.1

Getirana, A. C. V., Boone, A., Yamazaki, D., Decharme, B., Papa, F., & Mognard, N. (2012). The hydrological modeling and analysis platform (HyMAP): Evaluation in the Amazon basin. *Journal of Hydrometeorology*, *13*(6), 1641–1665. https://doi.org/10.1175/JHM-D-12-021.1

Getirana, A., Jung, H. C., Arsenault, K., Shukla, S., Kumar, S., Peters-Lidard, C., Maigari, I., & Mamane, B. (2020). Satellite Gravimetry Improves Seasonal

Streamflow Forecast Initialization in Africa. *Water Resources Research*, *56*(2), e2019WR026259. https://doi.org/10.1029/2019WR026259

Getirana, A., Kumar, S., Girotto, M., & Rodell, M. (2017). Rivers and Floodplains as Key Components of Global Terrestrial Water Storage Variability. *Geophysical Research Letters*, *44*(20), 10,359-10,368. https://doi.org/10.1002/2017GL074684

Getirana, A., Peters-Lidard, C., Rodell, M., & Bates, P. D. (2017). Trade-off between cost and accuracy in large-scale surface water dynamic modeling. *Water Resources Research*, *53*(6), 4942–4955. https://doi.org/10.1002/2017WR020519

Ghatak, D., Zaitchik, B., Hain, C., & Anderson, M. (2017). The role of local heating in the 2015 Indian Heat Wave. *Scientific Reports*, *7*(1), 7707. https://doi.org/10.1038/s41598-017-07956-5

Ghatak, D., Zaitchik, B., Kumar, S., Matin, M. A., Bajracharya, B., Hain, C., & Anderson, M. (2018a). Influence of Precipitation Forcing Uncertainty on Hydrological Simulations with the NASA South Asia Land Data Assimilation System. *Hydrology*, *5*(4), 57. https://doi.org/10.3390/hydrology5040057

Ghatak, D., Zaitchik, B., Kumar, S., Matin, M. A., Bajracharya, B., Hain, C., & Anderson, M. (2018b). Influence of Precipitation Forcing Uncertainty on Hydrological Simulations with the NASA South Asia Land Data Assimilation System. *Hydrology*, *5*(4), 57. https://doi.org/10.3390/hydrology5040057

Ghiggi, G., Humphrey, V., Seneviratne, S. I., & Gudmundsson, L. (2019). GRUN: An observation-based global gridded runoff dataset from 1902 to 2014. *Earth System Science Data*, *11*(4), 1655–1674. https://doi.org/10.5194/essd-11-1655-2019

Ghodichore, N., Dhanya, C. T., & Hendricks Franssen, H. J. (2022). Isolating the effects of land use land cover change and inter-decadal climate variations on the water and energy cycles over India, 1981–2010. *Journal of Hydrology*, *612*, 128267. https://doi.org/10.1016/J.JHYDROL.2022.128267

Gruber, A., Dorigo, W. A., Crow, W., & Wagner, W. (2017). Triple Collocation-Based Merging of Satellite Soil Moisture Retrievals. *IEEE Transactions on Geoscience and Remote Sensing*, *55*(12), 6780–6792. https://doi.org/10.1109/TGRS.2017.2734070

Gruber, A., Scanlon, T., van der Schalie, R., Wagner, W., & Dorigo, W. (2019). Evolution of the ESA CCI Soil Moisture climate data records and their underlying merging methodology. *Earth System Science Data*, *11*(2), 717–739. https://doi.org/10.5194/essd-11-717-2019

Gupta, P., Verma, S., Bhatla, R., Chandel, A. S., Singh, J., & Payra, S. (2020). Validation of Surface Temperature Derived From MERRA-2 Reanalysis Against IMD Gridded Data Set Over India. *Earth and Space Science*, *7*(1), e2019EA000910. https://doi.org/10.1029/2019EA000910

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., … Thépaut, J. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999–2049. https://doi.org/10.1002/qj.3803

Hu, Z., Liu, S., Zhong, G., Lin, H., & Zhou, Z. (2020). Modified Mann-Kendall trend test for hydrological time series under the scaling hypothesis and its application. *Hydrological Sciences Journal*, *65*(14), 2419–2438. https://doi.org/10.1080/02626667.2020.1810253

J. L. Monteith. (1965). *Evaporation and environment*.

Jacobs, C. M. J., Moors, E. J., ter Maat, H. W., Teuling, A. J., Balsamo, G., Bergaoui, K., Ettema, J., Lange, M., van den Hurk, B. J. J. M., Viterbo, P., & Wergen, W. (2008). Evaluation of European Land Data Assimilation System (ELDAS) products using in situ observations. *Tellus A: Dynamic Meteorology and Oceanography*, *60*(5), 1023–1037. https://doi.org/10.1111/j.1600-0870.2008.00351.x

Jin, X., Kumar, L., Li, Z., Feng, H., Xu, X., Yang, G., & Wang, J. (2018). A review of data assimilation of remote sensing and crop models. *European Journal of Agronomy*, *92*, 141–152. https://doi.org/10.1016/j.eja.2017.11.002

Kantha Rao, B., & Rakesh, V. (2019). Evaluation of WRF-simulated multilevel soil moisture, 2-m air temperature, and 2-m relative humidity against in situ observations in India. *Pure and Applied Geophysics*, *176*(4), 1807–1826. https://doi.org/10.1007/s00024-018-2022-7

Kim, N. H., & Office, N. D. (2017). *Global Soil Wetness Project Phase 3 Atmospheric Boundary Conditions (Experiment 1)*. https://doi.org/10.20783/DIAS.501

Kirchner, J. W. (2006). Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology. *Water Resources Research*, *42*(3). https://doi.org/10.1029/2005WR004362

Knoben, W. J. M., Freer, J. E., & Woods, R. A. (2019). Technical note: Inherent benchmark or not? Comparing Nash–Sutcliffe and Kling–Gupta efficiency scores. *Hydrology and Earth System Sciences*, *23*(10), 4323–4331. https://doi.org/10.5194/hess-23-4323-2019

Kumar, S., Peterslidard, C., Tian, Y., Houser, P., Geiger, J., Olden, S., Lighty, L., Eastman, J., Doty, B., & Dirmeyer, P. (2006). Land information system: An interoperable framework for high resolution land surface modeling. *Environmental Modelling & Software*, *21*(10), 1402–1415. https://doi.org/10.1016/j.envsoft.2005.07.004

Kumar, S. v, Peters-Lidard, C. D., Mocko, D., Reichle, R., Liu, Y., Arsenault, K. R., Xia, Y., Ek, M., Riggs, G., Livneh, B., & Cosh, M. (2014). Assimilation of Remotely Sensed Soil Moisture and Snow Depth Retrievals for Drought Estimation. *Journal of Hydrometeorology*, *15*(6), 2446–2469. https://doi.org/10.1175/JHM-D-13-0132.1

Landerer, F. W., & Swenson, S. C. (2012). Accuracy of scaled GRACE terrestrial water storage estimates. *Water Resources Research*, *48*(4). https://doi.org/10.1029/2011WR011453

Lohmann, D., Mitchell, K. E., Houser, P. R., Wood, E. F., Schaake, J. C., Robock, A., Cosgrove, B. A., Sheffield, J., Duan, Q., Luo, L., Higgins, R. W., Pinker, R. T., & Tarpley, J. D. (2004). Streamflow and water balance intercomparisons of four land surface models in the North American Land Data Assimilation System project. *Journal of Geophysical Research: Atmospheres*, *109*(D7). https://doi.org/10.1029/2003JD003517

Maina, F. Z., Kumar, S. V., Albergel, C., & Mahanama, S. P. (2022). Warming, increase in precipitation, and irrigation enhance greening in High Mountain Asia. *Communications Earth & Environment*, *3*(1), 43. https://doi.org/10.1038/s43247-022-00374-0

Maity, S., Satyanarayana, A. N. V., Mandal, M., & Nayak, S. (2017). Performance evaluation of land surface models and cumulus convection schemes in the simulation of Indian summer monsoon using a regional climate model.

Atmospheric Research, *197*, 21–41. https://doi.org/10.1016/J.ATMOSRES.2017.06.023

Mann, H. B. (1945). Nonparametric Tests Against Trend. *Econometrica*, *13*(3), 245. https://doi.org/10.2307/1907187

Mccollum, J. R., & Ferraro, R. R. (2005). *Microwave Rainfall Estimation over Coasts*. https://doi.org/10.1175/JTECH1732.1

McNally, A., Arsenault, K., Kumar, S., Shukla, S., Peterson, P., Wang, S., Funk, C., Peters-Lidard, C. D., & Verdin, J. P. (2017). A land data assimilation system for sub-Saharan Africa food and water security applications. *Scientific Data*, *4*(1), 170012. https://doi.org/10.1038/sdata.2017.12

Mizukami, N., Rakovec, O., Newman, A. J., Clark, M. P., Wood, A. W., Gupta, H. v, & Kumar, R. (2019). On the choice of calibration metrics for "high-flow" estimation using hydrologic models. *Hydrology and Earth System Sciences*, *23*(6), 2601–2614. https://doi.org/10.5194/hess-23-2601-2019

Nair, A. S., & Indu, J. (2019). Improvement of land surface model simulations over India via data assimilation of satellite-based soil moisture products. *Journal of Hydrology*, *573*, 406–421. https://doi.org/10.1016/J.JHYDROL.2019.03.088

Newman, A. J., Stone, A. G., Saharia, M., Holman, K. D., Addor, N., & Clark, M. P. (2021). Identifying sensitivities in flood frequency analyses using a stochastic hydrologic modeling system. *Hydrology and Earth System Sciences*, *25*(10), 5603–5621. https://doi.org/10.5194/hess-25-5603-2021

Niu, G. Y., Yang, Z. L., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., Kumar, A., Manning, K., Niyogi, D., Rosero, E., Tewari, M., & Xia, Y. (2011). The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements. *Journal of Geophysical Research Atmospheres*, *116*(12). https://doi.org/10.1029/2010JD015139

Niu, G.-Y., Yang, Z.-L., Dickinson, R. E., Gulden, L. E., & Su, H. (2007). Development of a simple groundwater model for use in climate models and evaluation with Gravity Recovery and Climate Experiment data. *Journal of Geophysical Research: Atmospheres*, *112*(D7). https://doi.org/10.1029/2006JD007522

Pai, D. S., Sridhar, L., Rajeevan, M., Sreejith, O. P., Satbhai, N. S., & Mukhopadhyay, B. (2014). *Development of a new high spatial resolution (0.25° × 0.25°) long period (1901-2010) daily gridded rainfall data set over India and its comparison with existing data sets over the region* (Vol. 65, Issue 1).

Patil, M. N., Waghmare, R. T., Halder, S., & Dharmaraj, T. (2011). Performance of Noah land surface model over the tropical semi-arid conditions in western India. *Atmospheric Research*, *99*(1), 85–96. https://doi.org/10.1016/J.ATMOSRES.2010.09.006

Rodell, M., Houser, P. R., Berg, A. A., & Famiglietti, J. S. (2005). Evaluation of 10 Methods for Initializing a Land Surface Model. *Journal of Hydrometeorology*, *6*(2), 146–155. https://doi.org/10.1175/JHM414.1

Rodell, M., Houser, P. R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C.-J., Arsenault, K., Cosgrove, B., Radakovich, J., Bosilovich, M., Entin, J. K., Walker, J. P., Lohmann, D., & Toll, D. (2004). The Global Land Data Assimilation System. *Bulletin of the American Meteorological Society*, *85*(3), 381–394. https://doi.org/10.1175/BAMS-85-3-381

Roy, P., Martha, T. R., Vinod Kumar, K., Chauhan, P., & Rao, V. V. (2023). Cluster landslides and associated damage in the Dima Hasao district of Assam, India due to heavy rainfall in May 2022. *Landslides*, *20*(1), 97–109. https://doi.org/10.1007/S10346-022-01977-6/METRICS

Running, S. W., Mu, Q., Zhao, M., & Moreno, A. (2019). *User's Guide MODIS Global Terrestrial Evapotranspiration (ET) Product (MOD16A2/A3 and Year-end Gap-filled MOD16A2GF/A3GF) NASA Earth Observing System MODIS Land Algorithm (For Collection 6)*.

Saharia, M., Jain, A., Baishya, R. R., Haobam, S., Sreejith, O. P., Pai, D. S., & Rafieeinasab, A. (2021). India flood inventory: creation of a multi-source national geospatial database to facilitate comprehensive flood research. *Natural Hazards*, *108*(1), 619–633. https://doi.org/10.1007/s11069-021-04698-6

Satish Kumar, K., AnandRaj, P., Sreelatha, K., & Sridhar, V. (2023). Reconstruction of GRACE terrestrial water storage anomalies using Multi-Layer Perceptrons for South Indian River basins. *Science of The Total Environment*, *857*, 159289. https://doi.org/10.1016/J.SCITOTENV.2022.159289

Sawada, Y., & Koike, T. (2016). Ecosystem resilience to the Millennium drought in southeast Australia (2001–2009). *Journal of Geophysical Research: Biogeosciences*, *121*(9), 2312–2327. https://doi.org/10.1002/2016JG003356

Shah, H. L., & Mishra, V. (2016). *Uncertainty and Bias in Satellite-Based Precipitation Estimates over Indian Subcontinental Basins: Implications for Real-Time Streamflow Simulation and Flood Prediction**. https://doi.org/10.1175/JHM-D-15

Shepard, D. (1968). A two-dimensional interpolation function for irregularly-spaced data. *1968*, 517–524.

Soni, A., & Syed, T. H. (2015). Diagnosing Land Water Storage Variations in Major Indian River Basins using GRACE observations. *Global and Planetary Change*, *133*, 263–271. https://doi.org/10.1016/j.gloplacha.2015.09.007

Srivastava, A., Sahoo, B., Raghuwanshi, N. S., & Singh, R. (2017). Evaluation of Variable-Infiltration Capacity Model and MODIS-Terra Satellite-Derived Grid-Scale Evapotranspiration Estimates in a River Basin with Tropical Monsoon-Type Climatology. *Journal of Irrigation and Drainage Engineering*, *143*(8). https://doi.org/10.1061/(asce)ir.1943-4774.0001199

Tellus. (2018). *TELLUS_GRACE_MASCON_CRI_GRID_RL06_V1*.

Upadhyaya, D. B., Evans, J., Muddu, S., Tomer, S. K., Bitar, A. Al, Yeggina, S., S, T., Morrison, R., Fry, M., Tripathi, S. N., Mujumdar, M., Goswami, M., Ganeshi, N., Nema, M. K., Jain, S. K., Angadi, S. S., & Yenagi, B. S. (2021). The Indian COSMOS Network (ICON): Validating L-Band Remote Sensing and Modelled Soil Moisture Data Products. *Remote Sensing 2021, Vol. 13, Page 537*, *13*(3), 537. https://doi.org/10.3390/RS13030537

Watkins, M. M., Wiese, D. N., Yuan, D. N., Boening, C., & Landerer, F. W. (2015). Improved methods for observing Earth's time variable mass distribution with GRACE using spherical cap mascons. *Journal of Geophysical Research: Solid Earth*, *120*(4), 2648–2671. https://doi.org/10.1002/2014JB011547

Xia, Y., Hao, Z., Shi, C., Li, Y., Meng, J., Xu, T., Wu, X., & Zhang, B. (2019). Regional and Global Land Data Assimilation Systems: Innovations, Challenges, and Prospects. *Journal of Meteorological Research*, *33*(2), 159–189. https://doi.org/10.1007/s13351-019-8172-4

Yamazaki, D., Ikeshima, D., Tawatari, R., Yamaguchi, T., O'Loughlin, F., Neal, J. C., Sampson, C. C., Kanae, S., & Bates, P. D. (2017). A high-accuracy map of global terrain elevations. *Geophysical Research Letters*, *44*(11), 5844–5853. https://doi.org/10.1002/2017GL072874

Yoon, Y., Kumar, S. V., Forman, B. A., Zaitchik, B. F., Kwon, Y., Qian, Y., Rupper, S., Maggioni, V., Houser, P., Kirschbaum, D., Richey, A., Arendt, A., Mocko, D., Jacob, J., Bhanja, S., & Mukherjee, A. (2019). Evaluating the Uncertainty of Terrestrial Water Budget Components Over High Mountain Asia. *Frontiers in Earth Science, 7*. https://doi.org/10.3389/feart.2019.00120

Yucel, I., Onen, A., Yilmaz, K. K., & Gochis, D. J. (2015). Calibration and evaluation of a flood forecasting system: Utility of numerical weather prediction model, data assimilation and satellite-based rainfall. *Journal of Hydrology, 523*, 49–66. https://doi.org/10.1016/j.jhydrol.2015.01.042

Zhang, X., Obringer, R., Wei, C., Chen, N., & Niyogi, D. (2017). Droughts in India from 1981 to 2013 and Implications to Wheat Production. *Scientific Reports*, *7*(1), 44552. https://doi.org/10.1038/srep44552

Zhao, W., & Li, A. (2015). A Review on Land Surface Processes Modelling over Complex Terrain. *Advances in Meteorology, 2015*, 1–17. https://doi.org/10.1155/2015/607181